SEQUENCE LISTING

<110> Fox, Michael
       Sullivan, John K.
       Holst, Paige
       Yoshinaga, Steven Kiyoshi

<120> B7-Like Polypeptides and Uses Thereof

<130> 00,759-A

<140>
<141>

<150> 60/233,867
<151> 2000-09-20

<160> 30

<170> PatentIn Ver. 2.0

<210> 1
<211> 1209
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (33)..(854)

<220>
<221> sig_peptide
<222> (33)..(89)

<220>
<221> misc_feature
<222> (693)..(755)
<223> predicted transmembrane domain

<400> 1
cagaaagaga cctatatgat caaatacaga ac atg atc ttc ctc ctg cta atg   53
                                    Met Ile Phe Leu Leu Leu Met
                                     1   5

ttg agc ctg gaa ttg cag ctt cac cag ata gca gct tta ttc aca gtg   101
Leu Ser Leu Glu Leu Gln Leu His Gln Ile Ala Ala Leu Phe Thr Val
         10   15   20

aca gtc cct aag gaa ctg tac ata ata gag cat ggc agc aat gtg acc   149
Thr Val Pro Lys Glu Leu Tyr Ile Ile Glu His Gly Ser Asn Val Thr
     25   30   35

ctg gaa tgc aac ttt gac act gga agt cat gtg aac ctt gga gca ata   197
Leu Glu Cys Asn Phe Asp Thr Gly Ser His Val Asn Leu Gly Ala Ile
 40   45   50   55

aca gcc agt ttg caa aag gtg gaa aat gat aca tcc cca cac cgt gaa   245

1

```
Thr Ala Ser Leu Gln Lys Val Glu Asn Asp Thr Ser Pro His Arg Glu
            60                  65                  70

aga gcc act ttg ctg gag gag cag ctg ccc cta ggg aag gcc tcg ttc   293
Arg Ala Thr Leu Leu Glu Glu Gln Leu Pro Leu Gly Lys Ala Ser Phe
            75                  80                  85

cac ata cct caa gtc caa gtg agg gac gaa gga cag tac caa tgc ata   341
His Ile Pro Gln Val Gln Val Arg Asp Glu Gly Gln Tyr Gln Cys Ile
            90                  95                  100

atc atc tat ggg gtc gcc tgg gac tac aag tac ctg act ctg aaa gtc   389
Ile Ile Tyr Gly Val Ala Trp Asp Tyr Lys Tyr Leu Thr Leu Lys Val
            105                 110                 115

aaa gct tcc tac agg aaa ata aac act cac atc cta aag gtt cca gaa   437
Lys Ala Ser Tyr Arg Lys Ile Asn Thr His Ile Leu Lys Val Pro Glu
120                 125                 130                 135

aca gat gag gta gag ctc acc tgc cag gct aca ggt tat cct ctg gca   485
Thr Asp Glu Val Glu Leu Thr Cys Gln Ala Thr Gly Tyr Pro Leu Ala
            140                 145                 150

gaa gta tcc tgg cca aac gtc agc gtt cct gcc aac acc agc cac tcc   533
Glu Val Ser Trp Pro Asn Val Ser Val Pro Ala Asn Thr Ser His Ser
            155                 160                 165

agg acc cct gaa ggc ctc tac cag gtc acc agt gtt ctg cgc cta aag   581
Arg Thr Pro Glu Gly Leu Tyr Gln Val Thr Ser Val Leu Arg Leu Lys
            170                 175                 180

cca ccc cct ggc aga aac ttc agc tgt gtg ttc tgg aat act cac gtg   629
Pro Pro Pro Gly Arg Asn Phe Ser Cys Val Phe Trp Asn Thr His Val
            185                 190                 195

agg gaa ctt act ttg gcc agc att gac ctt caa agt cag atg gaa ccc   677
Arg Glu Leu Thr Leu Ala Ser Ile Asp Leu Gln Ser Gln Met Glu Pro
200                 205                 210                 215

agg acc cat cca act tgg ctg ctt cac att ttc atc ccc tcc tgc atc   725
Arg Thr His Pro Thr Trp Leu Leu His Ile Phe Ile Pro Ser Cys Ile
            220                 225                 230

att gct ttc att ttc ata gcc aca gtg ata gcc cta aga aaa caa ctc   773
Ile Ala Phe Ile Phe Ile Ala Thr Val Ile Ala Leu Arg Lys Gln Leu
            235                 240                 245

tgt caa aag ctg tat tct tca aaa gac aca aca aaa aga cct gtc acc   821
Cys Gln Lys Leu Tyr Ser Ser Lys Asp Thr Thr Lys Arg Pro Val Thr
            250                 255                 260

aca aca aag agg gaa gtg aac agt gct atc tga acctgtggtc ttgggagcca   874
Thr Thr Lys Arg Glu Val Asn Ser Ala Ile
            265                 270

gggtgacctg atatgacatc taaagaagct tctggactct gaacaagaat tcggtggcct   934
```

                                    2

```
gcagagcttg ccatttgcac ttttcaaatg cctttggatg acccagcact ttaatctgaa 994

acctgcaaca agactagcca acacctggcc atgaaacttg ccccttcact gatctggact 1054

cacctctgga gcctatggct ttaagcaagc actactgcac tttacagaat taccccactg 1114

gatcctggac ccacagaatt ccttcaggat ccttcttgct gccagactga aagcaaaagg 1174

aattatttcc cctcaagttt tctaagtgat ttcca                          1209


<210> 2
<211> 273
<212> PRT
<213> Homo sapiens

<400> 2
Met Ile Phe Leu Leu Leu Met Leu Ser Leu Glu Leu Gln Leu His Gln
1               5                   10                  15

Ile Ala Ala Leu Phe Thr Val Thr Val Pro Lys Glu Leu Tyr Ile Ile
            20                  25                  30

Glu His Gly Ser Asn Val Thr Leu Glu Cys Asn Phe Asp Thr Gly Ser
            35                  40                  45

His Val Asn Leu Gly Ala Ile Thr Ala Ser Leu Gln Lys Val Glu Asn
        50                  55                  60

Asp Thr Ser Pro His Arg Glu Arg Ala Thr Leu Leu Glu Glu Gln Leu
65                  70                  75                  80

Pro Leu Gly Lys Ala Ser Phe His Ile Pro Gln Val Gln Val Arg Asp
                85                  90                  95

Glu Gly Gln Tyr Gln Cys Ile Ile Ile Tyr Gly Val Ala Trp Asp Tyr
            100                 105                 110

Lys Tyr Leu Thr Leu Lys Val Lys Ala Ser Tyr Arg Lys Ile Asn Thr
            115                 120                 125

His Ile Leu Lys Val Pro Glu Thr Asp Glu Val Glu Leu Thr Cys Gln
        130                 135                 140

Ala Thr Gly Tyr Pro Leu Ala Glu Val Ser Trp Pro Asn Val Ser Val
145                 150                 155                 160

Pro Ala Asn Thr Ser His Ser Arg Thr Pro Glu Gly Leu Tyr Gln Val
                165                 170                 175

Thr Ser Val Leu Arg Leu Lys Pro Pro Pro Gly Arg Asn Phe Ser Cys
            180                 185                 190

Val Phe Trp Asn Thr His Val Arg Glu Leu Thr Leu Ala Ser Ile Asp
            195                 200                 205

Leu Gln Ser Gln Met Glu Pro Arg Thr His Pro Thr Trp Leu Leu His
```

3

```
          210                  215                   220

Ile Phe Ile Pro Ser Cys Ile Ile Ala Phe Ile Phe Ile Ala Thr Val
225                     230                 235                 240

Ile Ala Leu Arg Lys Gln Leu Cys Gln Lys Leu Tyr Ser Ser Lys Asp
                245                 250                 255

Thr Thr Lys Arg Pro Val Thr Thr Thr Lys Arg Glu Val Asn Ser Ala
                260                 265                 270

Ile


<210> 3
<211> 254
<212> PRT
<213> Homo sapiens

<220>
<221> TRANSMEM
<222> (202)..(222)

<400> 3
Leu Phe Thr Val Thr Val Pro Lys Glu Leu Tyr Ile Ile Glu His Gly
1                   5                   10                  15

Ser Asn Val Thr Leu Glu Cys Asn Phe Asp Thr Gly Ser His Val Asn
            20                  25                  30

Leu Gly Ala Ile Thr Ala Ser Leu Gln Lys Val Glu Asn Asp Thr Ser
            35                  40                  45

Pro His Arg Glu Arg Ala Thr Leu Leu Glu Glu Gln Leu Pro Leu Gly
        50                  55                  60

Lys Ala Ser Phe His Ile Pro Gln Val Gln Val Arg Asp Glu Gly Gln
65                  70                  75                  80

Tyr Gln Cys Ile Ile Ile Tyr Gly Val Ala Trp Asp Tyr Lys Tyr Leu
                85                  90                  95

Thr Leu Lys Val Lys Ala Ser Tyr Arg Lys Ile Asn Thr His Ile Leu
            100                 105                 110

Lys Val Pro Glu Thr Asp Glu Val Glu Leu Thr Cys Gln Ala Thr Gly
            115                 120                 125

Tyr Pro Leu Ala Glu Val Ser Trp Pro Asn Val Ser Val Pro Ala Asn
        130                 135                 140

Thr Ser His Ser Arg Thr Pro Glu Gly Leu Tyr Gln Val Thr Ser Val
145                 150                 155                 160

Leu Arg Leu Lys Pro Pro Pro Gly Arg Asn Phe Ser Cys Val Phe Trp
            165                 170                 175
```

4

```
Asn Thr His Val Arg Glu Leu Thr Leu Ala Ser Ile Asp Leu Gln Ser
            180                     185                     190

Gln Met Glu Pro Arg Thr His Pro Thr Trp Leu Leu His Ile Phe Ile
            195                     200                     205

Pro Ser Cys Ile Ile Ala Phe Ile Phe Ile Ala Thr Val Ile Ala Leu
    210                     215                     220

Arg Lys Gln Leu Cys Gln Lys Leu Tyr Ser Ser Lys Asp Thr Thr Lys
225                     230                     235                     240

Arg Pro Val Thr Thr Thr Lys Arg Glu Val Asn Ser Ala Ile
                245                     250


<210> 4
<211> 224
<212> PRT
<213> Homo sapiens

<400> 4
Met Gly His Thr Arg Arg Gln Gly Thr Ser Pro Ser Lys Cys Pro Tyr
    1                   5                   10                      15

Leu Asn Phe Phe Gln Leu Leu Val Leu Ala Gly Leu Ser His Phe Cys
                20                      25                      30

Ser Gly Val Ile His Val Thr Lys Glu Val Lys Glu Val Ala Thr Leu
            35                      40                      45

Ser Cys Gly His Asn Val Ser Val Glu Glu Leu Ala Gln Thr Arg Ile
    50                      55                      60

Tyr Trp Gln Lys Glu Lys Lys Met Val Leu Thr Met Met Ser Gly Asp
    65                      70                      75                      80

Met Asn Ile Trp Pro Glu Tyr Lys Asn Arg Thr Ile Phe Asp Ile Thr
                85                      90                      95

Asn Asn Leu Ser Ile Val Ile Leu Ala Leu Arg Pro Ser Asp Glu Gly
            100                     105                     110

Thr Tyr Glu Cys Val Val Leu Lys Tyr Glu Lys Asp Ala Phe Lys Arg
            115                     120                     125

Glu His Leu Ala Glu Val Thr Leu Ser Val Lys Ala Asp Phe Pro Thr
    130                     135                     140

Pro Ser Ile Ser Asp Phe Glu Ile Pro Thr Ser Asn Ile Arg Arg Ile
145                     150                     155                     160

Ile Cys Ser Thr Ser Gly Gly Phe Pro Glu Pro His Leu Ser Trp Leu
                165                     170                     175

Glu Asn Gly Glu Glu Leu Asn Ala Ile Asn Thr Thr Val Ser Gln Asp
            180                     185                     190
```

5

Pro Glu Thr Glu Leu Tyr Ala Val Ser Ser Lys Leu Asp Phe Asn Met
          195   200   205

Thr Thr Asn His Ser Phe Met Cys Leu Ile Lys Tyr Gly His Leu Arg
      210   215   220

<210> 5
<211> 323
<212> PRT
<213> Homo sapiens

<400> 5
Met Gly Leu Ser Asn Ile Leu Phe Val Met Ala Phe Leu Leu Ser Gly
1   5   10   15

Ala Ala Pro Leu Lys Ile Gln Ala Tyr Phe Asn Glu Thr Ala Asp Leu
          20   25   30

Pro Cys Gln Phe Ala Asn Ser Gln Asn Gln Ser Leu Ser Glu Leu Val
          35   40   45

Val Phe Trp Gln Asp Gln Glu Asn Leu Val Leu Asn Glu Val Tyr Leu
      50   55   60

Gly Lys Glu Lys Phe Asp Ser Val His Ser Lys Tyr Met Gly Arg Thr
65   70   75   80

Ser Phe Asp Ser Asp Ser Trp Thr Leu Arg Leu His Asn Leu Gln Ile
              85   90   95

Lys Asp Lys Gly Leu Tyr Gln Cys Ile Ile His His Lys Lys Pro Thr
          100   105   110

Gly Met Ile Arg Ile His Gln Met Asn Ser Glu Leu Ser Val Leu Ala
          115   120   125

Asn Phe Ser Gln Pro Glu Ile Val Pro Ile Ser Asn Ile Thr Glu Asn
      130   135   140

Val Tyr Ile Asn Leu Thr Cys Ser Ser Ile His Gly Tyr Pro Glu Pro
145   150   155   160

Lys Lys Met Ser Val Leu Leu Arg Thr Lys Asn Ser Thr Ile Glu Tyr
              165   170   175

Asp Gly Ile Met Gln Lys Ser Gln Asp Asn Val Thr Glu Leu Tyr Asp
          180   185   190

Val Ser Ile Ser Leu Ser Val Ser Phe Pro Asp Val Thr Ser Asn Met
      195   200   205

Thr Ile Phe Cys Ile Leu Glu Thr Asp Lys Thr Arg Leu Leu Ser Ser

                                6

```
                210                    215                    220

Pro Phe Ser Ile Glu Leu Glu Asp Pro Gln Pro Pro Pro Asp His Ile
225                         230                 235                    240

Pro Trp Ile Thr Ala Val Leu Pro Thr Val Ile Ile Cys Val Met Val
                    245                 250                    255

Phe Cys Leu Ile Leu Trp Lys Trp Lys Lys Lys Lys Arg Pro Arg Asn
            260                 265                 270

Ser Tyr Lys Cys Gly Thr Asn Thr Met Glu Arg Glu Glu Ser Glu Gln
        275                 280                 285

Thr Lys Lys Arg Glu Lys Ile His Ile Pro Glu Arg Ser Asp Glu Ala
        290                 295                 300

Gln Arg Val Phe Lys Ser Ser Lys Thr Ser Ser Cys Asp Lys Ser Asp
305                 310                 315                    320

Thr Cys Phe


<210> 6
<211> 290
<212> PRT
<213> Homo sapiens

<400> 6
Met Arg Ile Phe Ala Val Phe Ile Phe Met Thr Tyr Trp His Leu Leu
  1               5                  10                  15

Asn Ala Phe Thr Val Thr Val Pro Asp Lys Leu Tyr Val Val Glu Tyr
                20                  25                  30

Gly Ser Asn Met Thr Ile Glu Cys Lys Phe Pro Val Glu Lys Gln Leu
            35                  40                  45

Asp Leu Ala Ala Leu Ile Val Tyr Trp Glu Met Glu Asp Lys Asn Ile
        50                  55                  60

Ile Gln Phe Val His Gly Glu Glu Asp Leu Lys Val Gln His Ser Ser
65                  70                  75                    80

Tyr Arg Gln Arg Ala Arg Leu Leu Lys Asp Gln Leu Ser Leu Gly Asn
                85                  90                  95

Ala Ala Leu Gln Ile Thr Asp Val Lys Leu Gln Asp Ala Gly Val Tyr
            100                 105                 110

Arg Cys Met Ile Ser Tyr Gly Gly Ala Asp Tyr Lys Arg Ile Thr Val
        115                 120                 125

Lys Val Asn Ala Pro Tyr Asn Lys Ile Asn Gln Arg Ile Leu Val Val
        130                 135                 140
```

7

```
Asp Pro Val Thr Ser Glu His Glu Leu Thr Cys Gln Ala Glu Gly Tyr
145                 150                 155                 160

Pro Lys Ala Glu Val Ile Trp Thr Ser Ser Asp His Gln Val Leu Ser
                165                 170                 175

Gly Lys Thr Thr Thr Thr Asn Ser Lys Arg Glu Glu Lys Leu Phe Asn
                180                 185                 190

Val Thr Ser Thr Leu Arg Ile Asn Thr Thr Thr Asn Glu Ile Phe Tyr
                195                 200                 205

Cys Thr Phe Arg Arg Leu Asp Pro Glu Glu Asn His Thr Ala Glu Leu
        210                 215                 220

Val Ile Pro Glu Leu Pro Leu Ala His Pro Pro Asn Glu Arg Thr His
225                 230                 235                 240

Leu Val Ile Leu Gly Ala Ile Leu Leu Cys Leu Gly Val Ala Leu Thr
                245                 250                 255

Phe Ile Phe Arg Leu Arg Lys Gly Arg Met Met Asp Val Lys Lys Cys
        260                 265                 270

Gly Ile Gln Asp Thr Asn Ser Lys Lys Gln Ser Asp Thr His Leu Glu
        275                 280                 285

Glu Thr
290


<210> 7
<211> 302
<212> PRT
<213> Homo sapiens

<400> 7
Met Arg Leu Gly Ser Pro Gly Leu Leu Phe Leu Leu Phe Ser Ser Leu
1                   5                   10                  15

Arg Ala Asp Thr Gln Glu Lys Glu Val Arg Ala Met Val Gly Ser Asp
                20                  25                  30

Val Glu Leu Ser Cys Ala Cys Pro Glu Gly Ser Arg Phe Asp Leu Asn
                35                  40                  45

Asp Val Tyr Val Tyr Trp Gln Thr Ser Glu Ser Lys Thr Val Val Thr
        50                  55                  60

Tyr His Ile Pro Gln Asn Ser Ser Leu Glu Asn Val Asp Ser Arg Tyr
65                  70                  75                  80

Arg Asn Arg Ala Leu Met Ser Pro Ala Gly Met Leu Arg Gly Asp Phe
                85                  90                  95

Ser Leu Arg Leu Phe Asn Val Thr Pro Gln Asp Glu Gln Lys Phe His
        100                 105                 110
```

8

```
Cys Leu Val Leu Ser Gln Ser Leu Gly Phe Gln Glu Val Leu Ser Val
        115                 120                 125

Glu Val Thr Leu His Val Ala Ala Asn Phe Ser Val Pro Val Val Ser
        130                 135                 140

Ala Pro His Ser Pro Ser Gln Asp Glu Leu Thr Phe Thr Cys Thr Ser
145                 150                 155                 160

Ile Asn Gly Tyr Pro Arg Pro Asn Val Tyr Trp Ile Asn Lys Thr Asp
                165                 170                 175

Asn Ser Leu Leu Asp Gln Ala Leu Gln Asn Asp Thr Val Phe Leu Asn
            180                 185                 190

Met Arg Gly Leu Tyr Asp Val Val Ser Val Leu Arg Ile Ala Arg Thr
        195                 200                 205

Pro Ser Val Asn Ile Gly Cys Cys Ile Glu Asn Val Leu Leu Gln Gln
        210                 215                 220

Asn Leu Thr Val Gly Ser Gln Thr Gly Asn Asp Ile Gly Glu Arg Asp
225                 230                 235                 240

Lys Ile Thr Glu Asn Pro Val Ser Thr Gly Glu Lys Asn Ala Ala Thr
                245                 250                 255

Trp Ser Ile Leu Ala Val Leu Cys Leu Leu Val Val Val Ala Val Ala
                260                 265                 270

Ile Gly Trp Val Cys Arg Asp Arg Cys Leu Gln His Ser Tyr Ala Gly
        275                 280                 285

Ala Trp Ala Val Ser Pro Glu Thr Glu Leu Thr Gly His Val
        290                 295                 300


<210> 8
<211> 316
<212> PRT
<213> Homo sapiens

<220>
<221> UNSURE
<222> (233)
<223> "Xaa" can be any naturally-occurring amino acid

<400> 8
Met Leu Arg Arg Arg Gly Ser Pro Gly Met Gly Val His Val Gly Ala
 1                   5                   10                  15

Ala Leu Gly Ala Leu Trp Phe Cys Leu Thr Gly Ala Leu Glu Val Gln
                20                  25                  30

Val Pro Glu Asp Pro Val Val Ala Leu Val Gly Thr Asp Ala Thr Leu
        35                  40                  45
```

9

```
Cys Cys Ser Phe Ser Pro Glu Pro Gly Phe Ser Leu Ala Gln Leu Asn
        50              55                  60

Leu Ile Trp Gln Leu Thr Asp Thr Lys Gln Leu Val His Ser Phe Ala
65                  70                  75                  80

Glu Gly Gln Asp Gln Gly Ser Ala Tyr Ala Asn Arg Thr Ala Leu Phe
                85                  90                  95

Pro Asp Leu Leu Ala Gln Gly Asn Ala Ser Leu Arg Leu Gln Arg Val
            100                 105                 110

Arg Val Ala Asp Glu Gly Ser Phe Thr Cys Phe Val Ser Ile Arg Asp
        115                 120                 125

Phe Gly Ser Ala Ala Val Ser Leu Gln Val Ala Ala Pro Tyr Ser Lys
            130                 135                 140

Pro Ser Met Thr Leu Glu Pro Asn Lys Asp Leu Arg Pro Gly Asp Thr
145                 150                 155                 160

Val Thr Ile Thr Cys Ser Ser Tyr Gln Gly Tyr Pro Glu Ala Glu Val
                165                 170                 175

Phe Trp Gln Asp Gly Gln Gly Val Pro Leu Thr Gly Asn Val Thr Thr
            180                 185                 190

Ser Gln Met Ala Asn Glu Gln Gly Leu Phe Asp Val His Ser Val Leu
            195                 200                 205

Arg Val Val Leu Gly Ala Asn Gly Thr Tyr Ser Cys Leu Val Arg Asn
        210                 215                 220

Pro Val Leu Gln Gln Asp Ala His Xaa Ser Val Thr Ile Thr Gly Gln
        225                 230                 235                 240

Pro Met Thr Phe Pro Pro Glu Ala Leu Trp Val Thr Val Gly Leu Ser
                245                 250                 255

Val Cys Leu Ile Ala Leu Leu Val Ala Leu Ala Phe Val Cys Trp Arg
        260                 265                 270

Lys Ile Lys Gln Ser Cys Glu Glu Glu Asn Ala Gly Ala Glu Asp Gln
        275                 280                 285

Asp Gly Glu Gly Glu Gly Ser Lys Thr Ala Leu Gln Pro Leu Lys His
        290                 295                 300

Ser Asp Ser Lys Glu Asp Asp Gly Gln Glu Ile Ala
305                 310                 315
```

<400> 9

```
Met Glu Ser Ala Ala Ala Leu His Phe Ser Arg Pro Ala Ser Leu Leu
 1               5                  10                  15

Leu Leu Leu Leu Ser Leu Cys Ala Leu Val Ser Ala Gln Phe Ile Val
            20                  25                  30

Val Gly Pro Thr Asp Pro Ile Leu Ala Thr Val Gly Glu Asn Thr Thr
            35                  40                  45

Leu Arg Cys His Leu Ser Pro Glu Lys Asn Ala Glu Asp Met Glu Val
         50                  55                  60

Arg Trp Phe Arg Ser Gln Phe Ser Pro Ala Val Phe Val Tyr Lys Gly
 65                  70                  75                  80

Gly Arg Glu Arg Thr Glu Glu Gln Met Glu Glu Tyr Arg Gly Arg Thr
                 85                  90                  95

Thr Phe Val Ser Lys Asp Ile Ser Arg Gly Ser Val Ala Leu Val Ile
            100                 105                 110

His Asn Ile Thr Ala Gln Glu Asn Gly Thr Tyr Arg Cys Tyr Phe Gln
         115                 120                 125

Glu Gly Arg Ser Tyr Asp Glu Ala Ile Leu His Leu Val Val Ala Gly
         130                 135                 140

Leu Gly Ser Lys Pro Leu Ile Ser Met Arg Gly His Glu Asp Gly Gly
145                 150                 155                 160

Ile Arg Leu Glu Cys Ile Ser Arg Gly Trp Tyr Pro Lys Pro Leu Thr
                 165                 170                 175

Val Trp Arg Asp Pro Tyr Gly Gly Val Ala Pro Ala Leu Lys Glu Val
            180                 185                 190

Ser Met Pro Asp Ala Asp Gly Leu Phe Met Val Thr Thr Ala Val Ile
         195                 200                 205

Ile Arg Asp Lys Ser Val Arg Asn Met Ser Cys Ser Ile Asn Asn Thr
210                 215                 220

Leu Leu Gly Gln Lys Lys Glu Ser Val Ile Phe Ile Pro Glu Ser Phe
225                 230                 235                 240

Met Pro Ser Val Ser Pro Cys Ala Val Ala Leu Pro Ile Ile Val Val
            245                 250                 255

Ile Leu Met Ile Pro Ile Ala Val Cys Ile Tyr Trp Ile Asn Lys Leu
            260                 265                 270

Gln Lys Glu Lys
         275
```

11

<210> 10
<211> 523
<212> PRT
<213> Homo sapiens

<400> 10

```
Met Glu Pro Ala Ala Ala Leu His Phe Ser Leu Pro Ala Ser Leu Leu
1               5                   10                  15

Leu Leu Leu Leu Leu Leu Leu Leu Ser Leu Cys Ala Leu Val Ser Ala
            20                  25                  30

Gln Phe Thr Val Val Gly Pro Ala Asn Pro Ile Leu Ala Met Val Gly
            35                  40                  45

Glu Asn Thr Thr Leu Arg Cys His Leu Ser Pro Glu Lys Asn Ala Glu
        50                  55                  60

Asp Met Glu Val Arg Trp Phe Arg Ser Gln Phe Ser Pro Ala Val Phe
65                  70                  75                  80

Val Tyr Lys Gly Gly Arg Glu Arg Thr Glu Glu Gln Met Glu Glu Tyr
                85                  90                  95

Arg Gly Arg Ile Thr Phe Val Ser Lys Asp Ile Asn Arg Gly Ser Val
            100                 105                 110

Ala Leu Val Ile His Asn Val Thr Ala Gln Glu Asn Gly Ile Tyr Arg
        115                 120                 125

Cys Tyr Phe Gln Glu Gly Arg Ser Tyr Asp Glu Ala Ile Leu Arg Leu
        130                 135                 140

Val Val Ala Gly Leu Gly Ser Lys Pro Leu Ile Glu Ile Lys Ala Gln
145                 150                 155                 160

Glu Asp Gly Ser Ile Trp Leu Glu Cys Ile Ser Gly Gly Trp Tyr Pro
                165                 170                 175

Glu Pro Leu Thr Val Trp Arg Asp Pro Tyr Gly Glu Val Val Pro Ala
            180                 185                 190

Leu Lys Glu Val Ser Ile Ala Asp Ala Asp Gly Leu Phe Met Val Thr
            195                 200                 205

Thr Ala Val Ile Ile Arg Asp Lys Tyr Val Arg Asn Met Ser Cys Ser
        210                 215                 220

Val Asn Asn Thr Leu Leu Gly Gln Glu Lys Glu Thr Val Ile Phe Ile
225                 230                 235                 240

Pro Glu Ser Phe Met Pro Ser Ala Ser Pro Trp Met Val Ala Leu Ala
                245                 250                 255

Val Ile Leu Thr Ala Ser Pro Trp Met Val Ser Met Thr Val Ile Leu
            260                 265                 270
```

                                    12

```
Ala Val Phe Ile Ile Phe Met Ala Val Ser Ile Cys Cys Ile Lys Lys
         275             280             285
Leu Gln Arg Glu Lys Lys Ile Leu Ser Gly Glu Lys Lys Val Glu Gln
         290             295             300
Glu Glu Lys Glu Ile Ala Gln Gln Leu Gln Glu Glu Leu Arg Trp Arg
305             310             315             320
Arg Thr Phe Leu His Ala Ala Asp Val Val Leu Asp Pro Asp Thr Ala
         325             330             335
His Pro Glu Leu Phe Leu Ser Glu Asp Arg Arg Ser Val Arg Arg Gly
         340             345             350
Pro Tyr Arg Gln Arg Val Pro Asp Asn Pro Glu Arg Phe Asp Ser Gln
         355             360             365
Pro Cys Val Leu Gly Trp Glu Ser Phe Ala Ser Gly Lys His Tyr Trp
         370             375             380
Glu Val Glu Val Glu Asn Val Met Val Trp Thr Val Gly Val Cys Arg
385             390             395             400
His Ser Val Glu Arg Lys Gly Glu Val Leu Leu Ile Pro Gln Asn Gly
         405             410             415
Phe Trp Thr Leu Glu Met Phe Gly Asn Gln Tyr Arg Ala Leu Ser Ser
         420             425             430
Pro Glu Arg Ile Leu Pro Leu Lys Glu Ser Leu Cys Arg Val Gly Val
         435             440             445
Phe Leu Asp Tyr Glu Ala Gly Asp Val Ser Phe Tyr Asn Met Arg Asp
         450             455             460
Arg Ser His Ile Tyr Thr Cys Pro Arg Ala Phe Thr Val Pro Val
465             470             475             480
Arg Pro Phe Phe Arg Leu Gly Ser Asp Asp Ser Pro Ile Phe Ile Cys
         485             490             495
Pro Ala Leu Thr Gly Ala Ser Gly Val Met Val Pro Glu Glu Gly Leu
         500             505             510
Lys Leu His Arg Val Gly Thr His Gln Ser Leu
         515             520


<210> 11
<211> 263
<212> PRT
<213> Homo sapiens

<400> 11
Phe His Val Ser Leu Leu Leu Val Gln Leu Leu Thr Pro Cys Ser Ala
 1               5               10              15


                              13
```

```
Gln Phe Ser Val Leu Gly Pro Ser Gly Pro Ile Leu Ala Met Val Gly
              20                  25                  30

Glu Asp Ala Asp Leu Pro Cys His Leu Phe Pro Thr Met Ser Ala Glu
          35                  40                  45

Thr Met Glu Leu Lys Trp Val Ser Ser Ser Leu Arg Gln Val Val Asn
      50                  55                  60

Val Tyr Ala Asp Gly Lys Glu Val Glu Asp Arg Gln Ser Ala Pro Tyr
65                  70                  75                  80

Arg Gly Arg Thr Ser Ile Leu Arg Asp Gly Ile Thr Ala Gly Lys Ala
                  85                  90                  95

Ala Leu Arg Ile His Asn Val Thr Ala Ser Asp Ser Gly Lys Tyr Leu
              100                 105                 110

Cys Tyr Phe Gln Asp Gly Asp Phe Tyr Glu Lys Ala Leu Val Glu Leu
          115                 120                 125

Lys Val Ala Ala Leu Gly Ser Asn Leu His Val Glu Val Lys Gly Tyr
          130                 135                 140

Glu Asp Gly Gly Ile His Leu Glu Cys Arg Ser Thr Gly Trp Tyr Pro
145                 150                 155                 160

Gln Pro Gln Ile Gln Trp Ser Asn Ala Lys Gly Glu Asn Ile Pro Ala
              165                 170                 175

Val Glu Ala Pro Val Val Ala Asp Gly Val Gly Leu Tyr Glu Val Ala
          180                 185                 190

Ala Ser Val Ile Met Arg Gly Gly Ser Gly Glu Gly Val Ser Cys Ile
          195                 200                 205

Ile Arg Asn Ser Leu Leu Gly Leu Glu Lys Thr Ala Ser Ile Ser Ile
      210                 215                 220

Ala Asp Pro Phe Phe Arg Ser Ala Gln Pro Trp Ile Ala Ala Leu Ala
225                 230                 235                 240

Gly Thr Leu Pro Ile Leu Leu Leu Leu Leu Ala Gly Ala Ser Tyr Phe
              245                 250                 255

Leu Trp Arg Gln Gln Lys Glu
              260
```

```
<210> 12
<211> 584
<212> PRT
<213> Homo sapiens

<400> 12
Met Lys Met Ala Ser Ser Leu Ala Phe Leu Leu Leu Asn Phe His Val
```

14

```
1                5                    10                   15
Ser Leu Phe Leu Val Gln Leu Leu Thr Pro Cys Ser Ala Gln Phe Ser
            20                  25                  30

Val Leu Gly Pro Ser Gly Pro Ile Leu Ala Met Val Gly Glu Asp Ala
            35                  40                  45

Asp Leu Pro Cys His Leu Phe Pro Thr Met Ser Ala Glu Thr Met Glu
        50                  55                  60

Leu Arg Trp Val Ser Ser Ser Leu Arg Gln Val Val Asn Val Tyr Ala
    65                  70                  75              80

Asp Gly Lys Glu Val Glu Asp Arg Gln Ser Ala Pro Tyr Arg Gly Arg
                85                  90                  95

Thr Ser Ile Leu Arg Asp Gly Ile Thr Ala Gly Lys Ala Ala Leu Arg
            100                 105                 110

Ile His Asn Val Thr Ala Ser Asp Ser Gly Lys Tyr Leu Cys Tyr Phe
        115                 120                 125

Gln Asp Gly Asp Phe Tyr Glu Lys Ala Leu Val Glu Leu Lys Val Ala
    130                 135                 140

Ala Leu Gly Ser Asp Leu His Ile Glu Val Lys Gly Tyr Glu Asp Gly
145                 150                 155                 160

Gly Ile His Leu Glu Cys Arg Ser Thr Gly Trp Tyr Pro Gln Pro Gln
                165                 170                 175

Ile Lys Trp Ser Asp Thr Lys Gly Glu Asn Ile Pro Ala Val Glu Ala
            180                 185                 190

Pro Val Val Ala Asp Gly Val Gly Leu Tyr Ala Val Ala Ala Ser Val
            195                 200                 205

Ile Met Arg Gly Ser Ser Gly Gly Gly Val Ser Cys Ile Ile Arg Asn
        210                 215                 220

Ser Leu Leu Gly Leu Glu Lys Thr Ala Ser Ile Ser Ile Ala Asp Pro
225                 230                 235                 240

Phe Phe Arg Ser Ala Gln Pro Trp Ile Ala Ala Leu Ala Gly Thr Leu
            245                 250                 255

Pro Ile Ser Leu Leu Leu Leu Ala Gly Ala Ser Tyr Phe Leu Trp Arg
        260                 265                 270

Gln Gln Lys Glu Lys Ile Ala Leu Ser Arg Glu Thr Glu Arg Glu Arg
    275                 280                 285

Glu Met Lys Glu Met Gly Tyr Ala Ala Thr Glu Gln Glu Ile Ser Leu
    290                 295                 300

Arg Glu Lys Leu Gln Glu Glu Leu Lys Trp Arg Lys Ile Gln Tyr Met
```

```
Ala Arg Gly Glu Lys Ser Leu Ala Tyr His Glu Trp Lys Met Ala Leu
                        325                 330                 335

Phe Lys Pro Ala Asp Val Ile Leu Asp Pro Asp Thr Ala Asn Ala Ile
                340                 345                 350

Leu Leu Val Ser Glu Asp Gln Arg Ser Val Gln Arg Ala Glu Glu Pro
            355                 360                 365

Arg Asp Leu Pro Asp Asn Pro Glu Arg Phe Glu Trp Arg Tyr Cys Val
    370                 375                 380

Leu Gly Cys Glu Asn Phe Thr Ser Gly Arg His Tyr Trp Glu Val Glu
385                 390                 395                 400

Val Gly Asp Arg Lys Glu Trp His Ile Gly Val Cys Ser Lys Asn Val
                405                 410                 415

Glu Arg Lys Lys Gly Trp Val Lys Met Thr Pro Glu Asn Gly Tyr Trp
                420                 425                 430

Thr Met Gly Leu Thr Asp Gly Asn Lys Tyr Arg Ala Leu Thr Glu Pro
            435                 440                 445

Arg Thr Asn Leu Lys Leu Pro Glu Pro Pro Arg Lys Val Gly Ile Phe
    450                 455                 460

Leu Asp Tyr Glu Thr Gly Glu Ile Ser Phe Tyr Asn Ala Thr Asp Gly
465                 470                 475                 480

Ser His Ile Tyr Thr Phe Pro His Ala Ser Phe Ser Glu Pro Leu Tyr
                485                 490                 495

Pro Val Phe Arg Ile Leu Thr Leu Glu Pro Thr Ala Leu Thr Ile Cys
                500                 505                 510

Pro Ile Pro Lys Glu Val Glu Ser Ser Pro Asp Pro Asp Leu Val Pro
                515                 520                 525

Asp His Ser Leu Glu Thr Pro Leu Thr Pro Gly Leu Ala Asn Glu Ser
        530                 535                 540

Gly Glu Pro Gln Ala Glu Val Thr Ser Leu Leu Leu Pro Ala His Pro
545                 550                 555                 560

Gly Ala Glu Val Ser Pro Ser Ala Thr Thr Asn Gln Asn His Lys Leu
                565                 570                 575

Gln Ala Arg Thr Glu Ala Leu Tyr
            580
```

<210> 13
<211> 526
<212> PRT

<213> Homo sapiens

<400> 13

```
Met Ala Val Phe Pro Ser Ser Gly Leu Pro Arg Cys Leu Leu Thr Leu
1               5                   10                  15

Ile Leu Leu Gln Leu Pro Lys Leu Asp Ser Ala Pro Phe Asp Val Ile
            20                  25                  30

Gly Pro Pro Glu Pro Ile Leu Ala Val Val Gly Glu Asp Ala Glu Leu
            35                  40                  45

Pro Cys Arg Leu Ser Pro Asn Ala Ser Ala Glu His Leu Glu Leu Arg
        50                  55                  60

Trp Phe Arg Lys Lys Val Ser Pro Ala Val Leu Val His Arg Asp Gly
65                  70                  75                  80

Arg Glu Gln Glu Ala Glu Gln Met Pro Glu Tyr Arg Gly Arg Ala Thr
                85                  90                  95

Leu Val Gln Asp Gly Ile Ala Lys Gly Arg Val Ala Leu Arg Ile Arg
            100                 105                 110

Gly Val Arg Val Ser Asp Asp Gly Glu Tyr Thr Cys Phe Phe Arg Glu
        115                 120                 125

Asp Gly Ser Tyr Glu Glu Ala Leu Val His Leu Lys Val Ala Ala Leu
        130                 135                 140

Gly Ser Asp Pro His Ile Ser Met Gln Val Gln Glu Asn Gly Glu Ile
145                 150                 155                 160

Cys Leu Glu Cys Thr Ser Val Gly Trp Tyr Pro Glu Pro Gln Val Gln
                165                 170                 175

Trp Arg Thr Ser Lys Gly Glu Lys Phe Pro Ser Thr Ser Glu Ser Arg
            180                 185                 190

Asn Pro Asp Glu Glu Gly Leu Phe Thr Val Ala Ala Ser Val Ile Ile
        195                 200                 205

Arg Asp Thr Ser Thr Lys Asn Val Ser Cys Tyr Ile Gln Asn Leu Leu
        210                 215                 220

Leu Gly Gln Glu Lys Lys Val Glu Ile Ser Ile Pro Ala Ser Ser Leu
225                 230                 235                 240

Pro Arg Leu Thr Pro Trp Ile Val Ala Val Ala Val Ile Leu Met Val
                245                 250                 255

Leu Gly Leu Leu Thr Ile Gly Ser Ile Phe Phe Thr Trp Arg Leu Tyr
            260                 265                 270

Asn Glu Arg Pro Arg Glu Arg Arg Asn Glu Phe Ser Ser Lys Glu Arg
        275                 280                 285
```

17

```
Leu Leu Glu Glu Leu Lys Trp Lys Lys Ala Thr Leu His Ala Val Asp
    290             295             300

Val Thr Leu Asp Pro Asp Thr Ala His Pro His Leu Phe Leu Tyr Glu
305             310             315                         320

Asp Ser Lys Ser Val Arg Leu Glu Asp Ser Arg Gln Lys Leu Pro Glu
                325             330                     335

Lys Thr Glu Arg Phe Asp Ser Trp Pro Cys Val Leu Gly Arg Glu Thr
                340             345             350

Phe Thr Ser Gly Arg His Tyr Trp Glu Val Glu Val Gly Asp Arg Thr
                355             360             365

Asp Trp Ala Ile Gly Val Cys Arg Glu Asn Val Met Lys Lys Gly Phe
370                     375             380

Asp Pro Met Thr Pro Glu Asn Gly Phe Trp Ala Val Glu Leu Tyr Gly
385                     390             395                 400

Asn Gly Tyr Trp Ala Leu Thr Pro Leu Arg Thr Pro Leu Pro Leu Ala
                405             410             415

Gly Pro Pro Arg Arg Val Gly Ile Phe Leu Asp Tyr Glu Ser Gly Asp
                420             425             430

Ile Ser Phe Tyr Asn Met Asn Asp Gly Ser Asp Ile Tyr Thr Phe Ser
            435             440             445

Asn Val Thr Phe Ser Gly Pro Leu Arg Pro Phe Phe Cys Leu Trp Ser
    450             455             460

Ser Gly Lys Lys Pro Leu Thr Ile Cys Pro Ile Ala Asp Gly Pro Glu
465             470             475                         480

Arg Val Thr Val Ile Ala Asn Ala Gln Asp Leu Ser Lys Glu Ile Pro
                485             490                     495

Leu Ser Pro Met Gly Glu Glu Ser Ala Pro Arg Asp Ala Asp Thr Leu
            500             505             510

His Ser Lys Leu Ile Pro Thr Gln Pro Ser Gln Gly Ala Pro
            515             520             525
```

<220>
<221> exon
<222> (4599)..(4652)
<223> coding portion of exon 1

<210> 14
<211> 7819
<212> DNA
<213> Homo sapiens

<220>

<221> unsure
<222> (145)..(292)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (1977)..(2459)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (2465)..(2576)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (2593)..(3410)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (3490)..(3533)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (6111)..(6194)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (6836)..(6870)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (7091)..(7112)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (7118)..(7420)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (7554)..(7575)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (7694)..(7718)
<223> "N" can be A, C, T, or G

<400> 14
ataagaagct gaattaaggt gatggcagtg gggtggaaga aaggagagcc accatgcaaa 60

```
aagtatccag gagggagaat taacaggact aggggatggg ccatatttgc aagatgagaa  120

atgcagaggt ctaagattct agctnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn  180

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn  240

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nngcataaac  300

catattttcc ccagaggagg attagtagga aaggaagctg ctggttggaa agtatcttta  360

tagcagtgtc tgttcctcgg tttgctcaag gggacagtgt gccaggaaag tccccgtgga  420

agggcaagga agaaggggaa gttaaagcca gtggcaggtg atccaagaat cttttctgtt  480

gctagagcta tgttacatgc tgtcctttca tgctctaaaa ataagagtgc tggcaagtgc  540

caggcctgtt ggtgcagctt aagatgatac ctttcttgga tatatatgca tctgaataag  600

gaaggctatc ttctggtcaa gctaaggtat gccatgagca tttccctgtg gaaagcactt  660

aattctgttc ccagttgtta cctgctgtaa gatctccctt tctaaaataa aaacaagaat  720

acagctcact gaggacctta catttccctc tagctactga ctcatttctc ttctcctttt  780

tatagcactc ttcttgagag agttgcctat atttgttgcc acatctttac ccattctctt  840

ttgaacctat tcaagctttc atctgtacaa aactcactga tactgtgctt gtcaggatca  900

tccatgacct ccatactgct aaatgcaact ctcaagagta tttggctcta ctgatcactc  960

ctttgtagca ctgtgtttta aaatataggt tttattatta tttaggtatg gtgaggccaa 1020

tatatcagga aatgactgtc gttgaaaaaa gtatgttgta ctcacagatc ccaagagaag 1080

ggggcacac catgccacaa agggccacat ggggaagcac caggtcagc caggaggtgg 1140

gtgggggtg cgcaagatct ttattgtggt ttcaacagga agaaatgggt gaagcagggt 1200

gagtggattt aggattagct gatataaata atttcagcag gctctggggc ataggggctg 1260

tccctagtct tctggtactt ggccctgggg tgattaaggc agttgcatag tgttgggaat 1320

gtgaaagccc ccaataaatg aggcagttgt gggtatggc tctgaaatgg gttggtttgc 1380

atttgaaagg tgtgctcatg ggcaagtggt ttactctctc ttagaggtta gaattggcta 1440

accctgggag cggcagtccc ttcagggtca gcaaggcccc aggtgtcaaa gcatcagaat 1500

acagaaaata aaatgcatgg ataatacaca ctgccatttg cctttgtacc cttcctttca 1560

atcttctctg ctggtgaccg ctcttcacaa agatctataa atgttggaat accccatgtc 1620

tcagtccttg ggcactctct ttcctatctc tctgtaggtg atgtaatgca gatatccatg 1680

actttaaatc tttaacactt ctgcattgat gactcctaaa tttacatctc taccccaact 1740

gcctactaaa cacctccact tggctatcta ataggcattt caaaccaaat ctacaacaaa 1800

                                  20
```

```
cgtaactctt tttccccttc cttaatttgc ttctccccca gccttctcca ttttaataaa 1860

cagcatctcc attgccttag tgactcaagc cccaaactta ggaattttcc cagatttccc 1920

tcttttctc aaactatata tctagcctgt cagcagttcc cttcaggtct tttttcnnnn 1980

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2040

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2100

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2160

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2220

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2280

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2340

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2400

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnna 2460

ctaannnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2520

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnngctc 2580

catttatatt tannnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2640

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2700

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2760

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2820

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2880

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2940

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3000

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3060

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3120

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3180

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3240

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3300

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3360

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn ttgtgtttta 3420

aatatatata cacacttaga cacatataac cctctttcgt atatcaatta tactttaata 3480
```

```
aagctgttgn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnctcgcca 3540

acagaccact tctcacccct actagcccct ctactctaag ctatcagcat ttctgcaaac 3600

acattcctaa ccgatctcac tgcttgtaat cttgccagca acctctccct ctcagcaata 3660

gtctattgcc tacaccaaag cttagttgtc tcttaatgat gtaaatgagg ttctatcatt 3720

ctcctgaccc aaaccctcca ctgcttttca tcacactcag agcagctctg ctgttgcctg 3780

atttagatgt atggctccaa cagatttccc ctgaagaaat gattccatgg ctgataaaag 3840

ttggaaagcc tcctcagttt cagaccatta tcagattagc tgtgtgctct gtccctttcc 3900

tcaaccataa gaagtccatg gataaagaaa gcttcagagt aaaggagaaa gcatgggagg 3960

tacagcagga ccaaggtggg gcattcgcag cccccaccct catcagagcc agttccctac 4020

tctccctgtc taaacctctt agtaagaggt agttcaagag agggcaaac tcaattccag 4080

cactcaaaag cacttgacta ctttgctcag tcaactagca agtatttatt gagaatgtag 4140

ctctgttcta tggagtctta ttttcaagtg tcagactccc agacatccag tccaggtaaa 4200

gaagatggtg tccattattc atttgacaaa caaagttggg gttcaagggc cagctattga 4260

aaaaagctat ggaaagcttc atgagacgtg caggtaactg ccaatatgtg tggttcacaa 4320

ggactggttc atattcagaa acggccatta gaaaaggaag aagaacttct catttggatt 4380

tataaagagt gtcttgttta ctcttaattt atatcttctc ttctccagga aatcaaccta 4440

taacttctcc tcccagctcc actctaccat ggtctgtcac cttccccaaa tgatttgtta 4500

ttccctgtt ttcaaaagtg aacaaagaac caaagaccca gcaaagtttc acaaggccct 4560

gagactttca attgtctatt tcagatcaaa tacagaac atg atc ttc ctc ctg cta 4616

atg ttg agc ctg gaa ttg cag ctt cac cag ata gca ggtaagaaag        4662

gacaaaggga gaggcttaag aaagaagagc aggtggtggt tcctagccaa agccaaaaat 4722

gagaatgtgg ccctcaggct gagggctttc tttgagagga cgtatgattt ctgggctatt 4782

ccaagcacca caaaaaaaaa aagagtcccc atggtggctt atacatgcca atgtccctat 4842

ctgacagaaa cggtgactga gaatattgct ccatctattc ccactatcca gtgagggtaa 4902

tgacaagaag acaggatcac tcagaccatg taaatctaaa ctgatacaag agggcagggg 4962

ttgagttccc ttaaaggtga gatgccaagc agctgtcccc ttcctttctg gcagggagag 5022

taaggagaca atggccaggg aacaccgtta ctctaaagat aatgtcttga agacattctg 5082

catattatta gttgtttctg tgagtttctt ttttgaaaag caacaatagc agccgttggt 5142

cattcatacc ttaatgtggt ttactgagtc ttcctaaaac ccaaatgaac aatgaacctt 5202
```

22

```
aaggctatcc ctttggactt gaagaaagga cttctattgg aggatgaggg tgagcagaaa 5262

gaaaagcagt ttcacagttg gttgttctcc tggggaaggt agttcagacc attcgagggt 5322

gtagttagaa ccatgagtgc actattttgg atgaacacca ggagctaaga gagtaacata 5382

gaggtgtgga cagaggatta agtcctcaag acaatagccc cagccccatg ggaaatcatc 5442

tttctgctca tgattgagaa ataatggctc ccttggcact tgataacctt tcgaagagct 5502

ttctcctccc tactagctgg ttccagatca ctcttcaccc agtcacattc ctctcactca 5562

cttgagctgc ccagcctggt ctggcactag agacatgcac ttggggccct cctcaaagga 5622

agaccctgag atattctgct tacttctact ctgctcctgc ctgcagggcc agctaaagga 5682

actttcatg ttttctttgc aaggaaccct gcctggctgg cattttagag acaagcaaaa 5742

gggggcaataa cttccttgct acaaaacagc ttcaagtttc catagagtga taagggaaat 5802

gagggccaaa agacactgtt ccccatcctg tggcaggact gggggcttca ggagaaaact 5862

tggggaatgt gtaacctctg tgggtttgta gcttaaaaac actgagatcc tgggtttct 5922

gtctttgttt tttgcctttt ctcttaggaa aggagtgagc tagggtgaca aggggcaaca 5982

tttttatcc ctcattggct ctttctacag aggaaggatc ttttcttcta agataatcag 6042

cacaagacaa tgaagatagg cactagctcc cagttaggta tactaatggg gcaaaaggaa 6102

gagcatttnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6162

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnattttgca ggtgagggaa ctagaactca 6222

gaaaaggtac ttaatttccc caagattaca tagttattag gtgacaccgg caagatttca 6282

acaaagctaa tgtcctttct actttactgt gctaccatga tgatggtaat caaaaatggc 6342

agacaaccca taaatcttcc aactttggaa taggtttttg cactgaagtc tgaatatgga 6402

tacgtattga atgtttattc tggatattca cagaatcaaa aaatatgtgt aatgaattat 6462

gttgctgaat taactgaaag gaaagtaaaa atgtagcgct ttctcatttt cttcacgaat 6522

ttggaattct tttctgcttt ccactatgca gataacatca gttcagacaa atattaaata 6582

cctacctaaa ttagaatgcc ttctcctcat gggatttttt taaaatcttg tcatttcatg 6642

tctctttaat taaagatgtt tgatttcaga ggagggtacc tgcaaaagaa aacaacaaaa 6702

aaactaaagg atctgagaaa taattagtgt ttacttctgg ggaggggagg aggtctggga 6762

tggggtaaa aaggatagtc ttatctatta tgtatattca ggtttttgtt ttttacaaga 6822

agcatgtatt aggnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnca agaaatttct 6882
```

23

```
catataaaaa tatgaaagta atcagactgc aacactcagt gcctgagaca gagctacagc 6942

tatcagggtg tccagacaga cagaagatta cattttcttc cttgctcctt gtacagcccc 7002

agacctgcat gcttcattga aaagaaaaga agatacctga attaaatcaa tgtgatgctt 7062

agtaccctat cagtgcacat ttcttttcnn nnnnnnnnnn nnnnnnnnnn cacttnnnnn 7122

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7182

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7242

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7302

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7362

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnca 7422

cttatggcgg tattctcagt cattacaaat aaataaaaac aatccatatg ccctggagaa 7482

tttgattcca ggagtaggtc tagaagaact tcaactggag aatggataga gaaatcatgg 7542

tatatttgca gnnnnnnnnn nnnnnnnnnn nnngatagca tgtgaataaa ttaattacaa 7602

aaacatatga ctacatctat tattatatag catgtagata aattacaaaa acatgtaact 7662

acatctatga atcttagagc ataatattga gnnnnnnnnn nnnnnnnnnn nnnnnnggcc 7722

agaagataac acatagcaca atgtcctttt cataaataaa tatattgctt aagcatacct 7782

tatatataga agataaagct taaaaagtaa agaagag                           7819
```

<210> 15
<211> 650
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> (20)..(175)
<223> "N" can be A, C, T, or G

<400> 15
```
attatcactt atgagggtgn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 60

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 120

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnctgcc 180

ctttatttt tcatgaaaga aattgctgaa gaggactaaa agaagtttta gtaagcattc 240

aataaatgta tgttctttat agtttccaaa tcagcaaata tagacatcct gcattttaa 300

ggagatttat atattttatt ggacatgctg taatttattt aaccacttcc ctgttggtag 360

acattatttc cattttcttc tgctagatta atgcttgaaa aaaatgtgtg cctcctaaag 420
```

```
actgtgatga aagttgcctc tgaataaaac tcaaacaaat cattaatcat taactctttc 480

cttacttgta tgctctttgg atgctctact gtgttatcta taaaataaag tttgaagtga 540

aaaattaggg taaaacattt tatatcattt ttaaaggata tatacatgga tgtacttaca 600

tatgcatgtt taaatttata taccataaca tttatttctt tttttaaaaa            650
```

<210> 16
<211> 9179
<212> DNA
<213> Homo sapiens

<220>
<221> exon
<222> (5211)..(5516)
<223> exon 2

<220>
<221> unsure
<222> (1054)..(1198)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (1297)..(1430)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (1482)..(1658)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (3286)..(3529)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (3580)..(3620)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (4193)..(4416)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (4506)..(4682)
<223> "N" can be A, C, T, or G

<220>
<221> unsure

<222> (6166)..(6223)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (6635)..(6803)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (7946)..(8351)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (8578)..(8609)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (8870)..(9178)
<223> "N" can be A, C, T, or G

<400> 16

```
tatatttgtt tttttctta catttttatt tcaaaatcta aggacatctt ataacccaga  60

aatattttt ataccttgtc atgtgttaga ggaaagagcc accccagtct tttttcattg  120

atgttttct tctctcttcg tactccagag gtagatgaaa accagagggc cacaatgacc  180

atggtgatgc ctgaggtcat tctgtgggcac agacctcagc ctaggttact ccacttcgcc  240

tatctttaga tccaaaacta ccctgctgac tgctgagata aacaaaggag aataatcagg  300

ttggggaaag gatttctatg cgaagacatg tctccatgca gtcctcctac actgagcaga  360

gcatgagtca ggtgcttaga gcaggatttt gtcctaaacc aggaacttca gagtttttctg  420

aagaatgtgg ctatgtaaag cacccccca ccccaccctt acttctcaag tacattacgt  480

ggcaagtctg aaaaaactta cacttctgtt gttaaatgtg ggggataaaa tataaactta  540

gtttcaagag gaagctatct tgggaggtaa tgcaaataat tcgttgtgtg tttcctgaat  600

aagtgacagg tgctgactac cattgatgct tcattgcaat aaaatgcaaa gctcccccaa  660

gaattttga aatgcatcaa gctaggtgtt ctaatctagc aaaaggacct gcatacatga  720

attttcatg cttttgccaa gtcttttgcc ctttagttta gttaagggcc ccacatgaat  780

ggaaagcctg tgttgtcagc ttaattttgt agttgtggaa accttccagt tttctccttt  840

gtctaatacc ttcaggagtt caatcctagg ttgaagctta atttaataac catgtggcat  900

gtaaagtaga aaacaaaaca tcttttcctt agcatacagc aaaaaaaaaa aaaaaactca  960

ctcatggatg tagtgtacac atgccagtgg atatatagtc ataactgcag tcattggtag 1020
```

26

```
cacagaaata aatgtgcatt gaagacacag agannnnnnn nnnnnnnnnn nnnnnnnnnn 1080

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1140

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnngg 1200

gctagcccag tgtctgactt actgtgttta agaaatatca actattacgc tacttcccag 1260

tgacagtcca aatgcagacc agtgttataa ctctacnnnn nnnnnnnnnn nnnnnnnnnn 1320

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1380

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn aggcaaggtg 1440

gtaaaagacg cttgcacgtg caaatgttac tttgtgtaac tnnnnnnnnn nnnnnnnnnn 1500

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1560

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1620

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnct tagtgaatct aatttgcagc 1680

tgggcttgag aaaaaacctc ttcatagaat tgtttgcatc agtgtcttga ttgcctctgt 1740

aacttacaat aagcaagaat gtttcaggat ttcaaaaatc tattgcattg cctaaacctc 1800

ttattttgta tggagtaatc aagctcaaag tttgcatgtc ttagaaactt tacttggggc 1860

aaaattagac caagtaacaa ttaatcttct aggtattctg agctattcag acatatgatt 1920

catgtttgct aattgctctt ttctcttgta aatattagct gaaaaatgtc acctgtctga 1980

caagtagcat attttatgcc tatcactcct ggcacgcatt cttacaaggc agacaggaaa 2040

aataggaaga aaatggactt ttatcaaagg cccaggcagt aaagagggga gttctgctgt 2100

aagctaaggg gagttccaga ggaagttata ggcgttccct ttcttatgac aagaaagcat 2160

agtgcagtaa ataaatttgc taaatagatt caacagtctc tacccaaagt catctattta 2220

attcttgttg ttatgcagac tcagcaacta accttccttg taagccccat tttcttccct 2280

gtttcctgtt tatcaaatgt aattaaacaa gagaagtatt atagaagagt aaaagtagta 2340

ggtaattctt gaacttggca tatgattact acatatttga tgaatagttg aatattattc 2400

ttcaaggaca gattggattt ggtatcaggt ggctctgcat taagttataa gggacttaat 2460

aactcaagta tttaaggacg gcttccatca taaagggatc tgcccttaag agggtcccat 2520

tatggagatt ctgaggtgag agctattcca agtgtgcagt ggattaaaat aaaagaatca 2580

tacaggaaat ctctttttac atgccttatt ccagggtctt tgcaacctgg cacagcaagt 2640

gcagatatga ttagcattgt tttacacatg tacactcacc ttatagccct gcccctgtgc 2700
```

27

```
ccctcctgca caaaagaatg ctgggcacac gtgaactcct ctctgtagaa aggcacatta 2760

atgttctagc catggttaaa acaggggatag aggcaagcca aaaatgtcgg tcatttgaaa 2820

taaatctcaa gtttgtgcat atcactatca agtgtgctgt gtggcaatta agaatgccaa 2880

tttgtgtgat cacaggcaag ttgcagtttg atgaaaggaa agcagaggtg aatatataac 2940

cagggtcatc ctttctttct ccctctctct ctttctgtca tttatttgcc aagctcttaa 3000

ctagaacttg ctatgtgcta ggtactggat atatcaaagc aaactcagcc tggtctttgc 3060

cttcaaagat ttgcaggata gtgggaagaa aaacttgaat cagaggacat ctgcagtggg 3120

aatcattcaa gcacagaaa acccaaaagt tacttatact gtgaaatctg atcagagaat 3180

ggactgtcct ggttagtaaa atatcctgga ggataaagat tggccatgca ttccacatat 3240

gaattaccac tttcccaaga attaaaacat ggtacgaaag aaaggnnnnn nnnnnnnnnn 3300

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3360

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3420

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3480

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnna aacaacttct 3540

ggaatagcta atgcttagaa gcagctccca aatatttgtn nnnnnnnnn nnnnnnnnnn 3600

nnnnnnnnnn nnnnnnnnnn gcaagctcta ctgaacataa tttgatctaa tcttctgtga 3660

ttattcagaa actacttcaa gattttccta tacctccatc ataatgaata cccattcatt 3720

aatgatggaa gcagcctaat tttgtcattt ttcacacttt attgatgtaa cactaccttt 3780

actagtttgg ccactcctta tgcttttttt atagaactat ttagatcaat tcaactttta 3840

aaaaataaag ccacataccc ctgtggtaga tgaaaaacaa gtatcatttg cactggtaaa 3900

tagagaatag gaagaaaaat aaatgcagtg aaaataaagc agtgttatca aatcctaccc 3960

agatactgtt atctacccgg aagcttcctg tttgattaaa aggaaaaata gccagtgtta 4020

gaggtgtgga agtctagttg aaattatatg caattgaagg attaaaatag aattgaaaag 4080

ggaataaatt cctctctgaa taatttaact ccctttaggc tttgattctg cctcatctaa 4140

aatcatctta catacttcta gtggcgtgtc cctcacattt tggtaaactc tgnnnnnnnn 4200

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4260

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4320

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4380

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnccat ttttcctttc cctttattgt 4440
```

28

```
cagaaaatag aaagcatcta cagtgggctt gtatgatgtg gtggttagaa atacctgatc 4500

tgattnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4560

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4620

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4680

nnaccatcac tggtgagagg aagtgatacc tggcacaaaa atatatggat taatcaatat 4740

ggattgaggg aaacaaacct ggagaatagg atgtgaaggt atttaagtaa catgagctca 4800

gaccttgatg gtagggaagt cgaaaggaag cattttgttc ttatatgaca gatgacctgg 4860

aatgactgca gggcttgggg ggtcagggac tggaggtggg agaggcctct gagagcaagc 4920

agtgctgtcc accagaagct cttgctgggg tgcccagaga ggagcaaagg gcagtcagct 4980

gcacaggagg gaatgtttgg aggagagagc cacctcagat cagcgggtca agaatccac 5040

tcttgcccag atggatgggg caaaggagaa aaaggattcg ccacgggaat gtccagataa 5100

gacaggtgcc ttttggaaaa tggggtgag atgggtctca ggttacactt cgtaagaact 5160

ggaatgtaaa gtaaaggcag acaatgacaa aatatcttgt tttcttttca gct tta     5216

ttc aca gtg aca gtc cct aag gaa ctg tac ata ata gag cat ggc agc  5264

aat gtg acc ctg gaa tgc aac ttt gac act gga agt cat gtg aac ctt  5312

gga gca ata aca gcc agt ttg caa aag gtg gaa aat gat aca tcc cca  5360

cac cgt gaa aga gcc act ttg ctg gag gag cag ctg ccc cta ggg aag  5408

gcc tcg ttc cac ata cct caa gtc caa gtg agg gac gaa gga cag tac  5456

caa tgc ata atc atc tat ggg gtc gcc tgg gac tac aag tac ctg act  5504

ctg aaa gtc aaa ggtgagtggt gtcaaggact agaatccatg gaagctctct       5556

ccaacagagg atctgcaagt cacagaaacc cattaaaggt agctcaagca aaaacaagca 5616

ggctgctttt aaggagacag ctatttcaga gaaaatgaaa gcatctgctc ggaaataatt 5676

tttgacatct gagtacaaag cagccgaagt acaagtgaaa gggggtagga cctataggaa 5736

taaaatggga ctggaggaag ccaggaaaat tagtccctga aatgtgggag ggtatgaaaa 5796

ataagctttg cctaattcac aattctccca tggaacatcc ctgacttgat tattaagata 5856

ctcttttca atagtttata ccctgaatcc agagtttta aaaccatggt ttgccgccca 5916

ttcatggatt aaaatatcaa tttagtgagt agcaaccaga tgcacgtttc ccgcccttta 5976

aaaaataatg tatagaagag aatagacaga gtagatcaga cgatatcaca gagtaggact 6036
```

```
gagtactgta aaactaattt ctgagggacg tgtgtgtgtg tgtgcgtgtt gggtcatggt 6096

ataaattttt ttttttcttac tttggatcat aaaaagttac aagtttggaa aacactgctc 6156

aaatgcaagn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6216

nnnnnnnaga gtcccatgaa gtctatagct gtccctattc ctctggattc agggatctct 6276

ccactccagc acaattgaaa atctaaatat aaagagaatc ttcacactct tgtttgttct 6336

agaaaaggtg atttgaggaa agacatataa caactataaa aaatagattt tgcttgttca 6396

ttggcttatg gtctccaggc ttgaatgctc tgagataaat gatgccaata tttctctggc 6456

ctcttcccat cccacgcatt ggacctcaga tggtctgtac tgtcttctag agggtttgtg 6516

ggttttggcc ccaaaaaacc attaaccttg gcagaaagtg tgtgacttta tgatctggta 6576

caaagaagga caaactagag ggactggaca tgaggatgaa tattgtgttc gcccttatnn 6636

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6696

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6756

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnaaa gaccctcagt 6816

tgttacaggg gcagtgacct cctcacacct caaccatcaa tgagtcacca ggaaagccat 6876

tagcctagat gtaactgttt tctatcttta ttgcatttcc tacatccagg cagcagctgg 6936

gaggaactct agaacactga agtttgtctg agttccctta atgtaaggct gtacattctc 6996

aggatgcctt gatgtactcg aatatctgca accctaaatc accacctctg tttttattga 7056

tctctatctg aatgctgtat taatgggcca ggccttctgc ccattctctc aaactgagaa 7116

ctgtctctca ttcctgggga ggcaccctgc ctactcctta cctagatcag ggatttctca 7176

gttgtggaga gatttgttcc ttatagtgtt ggtcatcaaa ctgggatatt tggggattac 7236

aaagactttt caagggatgt atgggcacag gcagtttag gaagtgagtt cctagatcct 7296

catctcccc aaatactcgt tcccaaaatt gacgagcctg acaatgtgca tgccaggcaa 7356

ggctcttggg gttcccctaa aacacttcct cttttaagcc taccactcac tcatcatgaa 7416

tatagtccat tgtcccaggg tgtaaaaccc tctatagtgt taaataaaag aatgattggg 7476

aacattgaca cctgatggaa ctgttatgac taaaaaccct tttgcaaata atgtggtatc 7536

taattttctg ctttcaacaa aattgaagga ggcccttata aagttaataa ctgataatca 7596

aaaatgagta atttttgcca tgtaaatcag gctaaagaat gaaatggcat tgctgtaacg 7656

aaactgcttc cattcccatt gatttactca tacgaacaag attccttagc ctttataagc 7716

tacaaaaaaa tgaaaaatag aaatagaatt gaggctgaat tctattatat aaaatcattc 7776
```

                                    30

```
caaccatgtc atatggttct tcggattcat gaataatttg gaaaagagag ccatatccat 7836

cttattaagg gacacattcc caataaattt tcatctttca tgtttaataa ttatcaatat 7896

tcataacatt ttacattttg atcaaatatg tgtttaataat aatagaaatn nnnnnnnnnn 7956

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8016

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8076

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8136

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8196

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8256

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8316

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnattag atcttaatgc agaacaccct 8376

gaacatttaa agcttcatag tcacaagaga aaagttttca tttcaatagc tataaatatt 8436

ttgttgttgt aaagacatat aacgataatc aatacaaaat ctgtcaaaca aaaatatgtt 8496

acattaagat aaaattctgt agggaaggtg aaattggaag tgagtttcaa tgaatgaaaa 8556

gaaacaattt agacagagaa gnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnncttagaa 8616

ggaattgaat agattaggtt ccttacccaa aaagcctctg ttatttgtct tatttattta 8676

ttctcttttt tccacattcc ccagtctcat tcccctttt taacacagga aattattcca 8736

gcatgtttca tacatattct tttgtttgta agagcttatt taaaatatgt aatattgttt 8796

tagatgcata tattttttt cttgttggaa ctatattgta ctatatatat atattttaga 8856

aatggacaca ttannnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8916

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8976

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9036

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9096

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9156

nnnnnnnnnn nnnnnnnnnn nnt                                       9179
```

```
<210> 17
<211> 1814
<212> DNA
<213> Homo sapiens

<220>
<221> exon
```

<222> (738)..(1010)
<223> exon 3

<220>
<221> unsure
<222> (72)..(118)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (405)..(630)
<223> "N" can be A, C, T, or G

<400> 17

```
tagatctcag ctttcttgag gcagggagcc atatctgttt aattcactca gcatatactg   60

caaagaagca gnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnta  120

gggagaagtg cagaataaat atacccaact ctttactatg tatagacatt atctaggtct  180

ttattttttt tctcttctta atctcaaaga aaacagagga aaggaggaag taaaaagtaa  240

atttttgcct gaagatgttt ggaaaaaata ccaaataaag tgagatagtg ggtaatctag  300

tgatttttat ttttccgtcc tctttctggc ctccaattgt gaaataattt atagcactgt  360

aagaaagaag ccacaaattg tggtagcttg gaccactgtt gaggnnnnnn nnnnnnnnnn  420

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn  480

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn  540

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn  600

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn acttaatatt gatagtgata attttattca  660

tttcacatgg catgagtac caagctctat aggaatcaga aaataaagtc ttatttctttt  720

ttcttctcta ttgtcca gct tcc tac agg aaa ata aac act cac atc cta      770

aag gtt cca gaa aca gat gag gta gag ctc acc tgc cag gct aca ggt     818

tat cct ctg gca gaa gta tcc tgg cca aac gtc agc gtt cct gcc aac     866

acc agc cac tcc agg acc cct gaa ggc ctc tac cag gtc acc agt gtt     914

ctg cgc cta aag cca ccc cct ggc aga aac ttc agc tgt gtg ttc tgg     962

aat act cac gtg agg gaa ctt act ttg gcc agc att gac ctt caa agt    1010

aagagctgcc cccacttcct aggtctatca gttagggttc agacaagaaa cagatgtgcat 1070

actcgagtga tttgaggaga gtgtaataaa gggactgttt acaaaggtgt gatcaccatt 1130

tggagaaact acaaaggata gtgcagaaca ctggggcttc aatgttggga gggcaattac 1190

cactgttgga gaagttactg gaatcagaag ggagctgtag ggaaagcccc acttcccagg 1250
```

32

```
agctgtagcc acagaatagg gaagctgcca catgcagcga ctccaaaggg tggaaactgg 1310

atgaatgaat accccaactc attctcctcc caccctccaa tctcctgcta gcacctccca 1370

ttggctgaac ccagctagaa gtcagagaat acaagggtcc actgttgtat tccataaaag 1430

tcaacttctc agggctcaga gcaatattga catgtacaga atagatctgg agaggaaaca 1490

gaaaatatct agtacaatag ctaatcactg tgattcatgc acagtgtcat gagccagcag 1550

gatgaatatt cctttgctgt acttgctgcc agtcagctgg ttatgggttt ttccaagaaa 1610

tttggtctct aacaaaattc ttcagagcct ttactgacta tgctggatat ttttggaagg 1670

gatcccatac ttttgaactt catacagcag aatttcaaac aatcttggga aaataacaac 1730

ttttatctgc ccagtaagga caactaacac ctagtatcat aatcatttcg taagagacag 1790

gtaatttcat caccgagtgc atat                                        1814
```

```
<220>
<221> unsure
<222> (3929)..(4224)
<223> "N" can be A, C, T, or G


<220>
<221> unsure
<222> (5233)..(5352)
<223> "N" can be A, C, T, or G


<220>
<221> unsure
<222> (5974)..(6030)
<223> "N" can be A, C, T, or G


<220>
<221> unsure
<222> (6048)..(6073)
<223> "N" can be A, C, T, or G


<220>
<221> unsure
<222> (6091)..(6306)
<223> "N" can be A, C, T, or G


<220>
<221> unsure
<222> (7576)..(7688)
<223> "N" can be A, C, T, or G


<220>
<221> unsure
<222> (8259)..(8368)
<223> "N" can be A, C, T, or G


<220>
<221> unsure
<222> (9787)..(9885)
<223> "N" can be A, C, T, or G


<220>
<221> unsure
<222> (9926)..(10110)
<223> "N" can be A, C, T, or G


<220>
<221> unsure
<222> (10113)..(10241)
<223> "N" can be A, C, T, or G


<220>
<221> unsure
<222> (11000)..(11438)
<223> "N" can be A, C, T, or G


<220>
<221> unsure
```

<222> (12528)..(12844)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (13720)..(13963)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (13983)..(14275)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (14529)..(14551)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (14585)..(14621)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (14703)..(14749)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (16380)..(16436)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (16484)..(16517)
<223> "N" can be A, C, T, or G

<220>
<221> unsure
<222> (17022)..(17366)
<223> "N" can be A, C, T, or G

<400> 18
ctactgagaa gggatatact ctcagactaa aggacagtcc ctagtactga ttcaatctgg 60

ctttatagaa aattcactat attgtcattg tatttcacag tttgcccttt gtcttagctg 120

gtaagacaga gcctatgata aggacttgtg tggcatgcag gtatttaatt ggcaacccca 180

gagggcagaa gcaagagatt taggagttta agagagggta atataagagt atattatcaa 240

agttgtagtg tggacaacag aaaactcaaat attcaaggac cagcatgtag acagcctcct 300

aagatgtcta ctcagacaaa gaatttcagg tggaaggact tgttcatctg cttcacgccc 360

attggttgac aggaatatga actccattct gctgctgggc tagacatgca tgtgggctga 420

```
gtgagctttc cccagtatcc gtagcatcag aaaagtcgca gggcagaaag aaaagtatcc 480

aatttgaggt gaattactga ccttgaagtg agtgtaagcc taactagaat tctaccccag 540

ctggctgaag tgaaaggtga ggctgagagg aaataaggca ggactgcaca gtccccaatt 600

gtactgttca aatccactca tgcccttcat taagtcagct ctgccactga gccttccagc 660

tgggaggcag ccacaatctc tgcagaagat ttaatataca ccagtttgtg gaacaagctg 720

tnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 780

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 840

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 900

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 960

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1020

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1080

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1140

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1200

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1260

nnnnnnnnnn nnnnnnnnacc tgcagcactg ggatagccct ggtacagacg tagaacatgc 1320

ttgaggaggt tgtcaggagg aaatgagtta gaccttgcac agaactacca ccatcaagta 1380

cagttggggt gaggcagaca tggtatgagt tgaggcatca gagatatctg atgcttttatg 1440

ccaaattaaa attaattttt tcatggagtg acactgatcc acagaccaga ctccaagaac 1500

tttgcagtga ctaaataccc atctcatcat aacttttcctg gtattttctt ctggaaaaaa 1560

ttcttccctg atacagtttt cagaggcagc tagatgcact gtcatctctc ccctttttccc 1620

acttccctac ctatccacaa tttactaccc aatgccaaca ctaaagttag cccaacttcc 1680

ttctaactaa attattagtt tagaaggaaa gagaggagtc atgctaagga tcttaactgn 1740

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1800

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1860

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1920

nnnngggggag aggaggggag aggaggtagg gagggggaaag aaagaaacta gaaatccatc 1980

aattttagga ccaacttcag gtaaaaaaat gaattaggca agttggtctt tcaacattct 2040

ctacctctct ttatatcatg gttgagacca cagacttctc acctcatgaa agatgaactc 2100
```

```
taactaattc atactaaagc taaagcctct aaagaggatt aaatatgagc aatcccacga 2160

gaactttttt cccctggaat tgtttattca actgtcgttc gttatatgga atttcctgcc 2220

tggttaagtg taggccagta ctttggatga attgtagttt tctagaaaga cgcttcttat 2280

ataagaacct ctccaggaaa acaggggcct gtatgagatg aattgagaaa taactttaca 2340

ccactgatta tgtcagtgtt ctattctgca tggtagagat gtgaaagggc agactgacca 2400

ttgctctgga agcctttacg ctgtgagaag ttaacagtag agtaaaatgg ccactccact 2460

ctcttcatgg aagccaacat ggcttactaa atagtcaaca accatgggag agacctgtgg 2520

ggtcttcatc agagctcagg atctcctagg gtatcactca taaatacagc catcagggag 2580

atggagaaat ctttgtgcag ccagaaattc tcaacctggt tttacccatc cttcccaact 2640

ttgtattcgt cctactgttt actgacatgg atcctctgct tcattaacca tcccttcctc 2700

accacatgct ctctgaactt ggctgcacct tttctacctc catgccttct ttgctcaggt 2760

ttttccacat aaatatcatt atttccctct ctactagctc caagcccacc ctctctctgg 2820

ggcagctcag tcactccagg gcacaagggg gtctttccct catcccacat tttgagacct 2880

actacctgga ccatttgttt gccttgtaac tatgcttgcc ttttttaattg ctattttatt 2940

ttccatgtat tttcattgtt cacacaagtc ttctttattc cacactaagg caaaagcaga 3000

gtcctgtgtt cataataagt gctcaacaaa tgttgggttg attggggttgg agattccatc 3060

ttagataatc gcagtcccat catgccagct accagactgt gtggacagcc aggtcagagc 3120

agccaaatga tattctagct tgtggcacaa ataccagcaa caaaataacc aaagtcacac 3180

atctgcctct gagttcctgg cttctatttc tcaagggcat ttttaagttg tcttatgact 3240

gttcccttc tactcattct cataaattga gctgtggact gctgtgaccc acaagcttct 3300

ccggaagtca atgtataaaa caaacacgga aacgaagagt atggtgggtg gagggtactc 3360

cactgactct agaatggatg actgaacatt ccaaatttca agcacaagtt agggagcaac 3420

agatcatttt ccttttgaaa tagggtttct tctgctcagc cagttgttgt attttcatta 3480

ggaaatggaa tgggactaca gcacaaaaaa taaatataaa aggacccttg tagggctggc 3540

agaaaagaga atccttccta ggagacctgg aggtgattcc aggcnnnnnn nnnnnnnnnn 3600

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3660

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3720

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3780

nnntcagaaa gtgtgcaaac agtaaaaaaa aatggtatat ctagcaagtt gcatgcctta 3840
```

```
cttgtgagtt catgaagttg tggcaaggat aagacaaata tttttgcca ttgcatcatt 3900

atatcattgc taagagtatg ccattattnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3960

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4020

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4080

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4140

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4200

nnnnnnnnnn nnnnnnnnnn nnnnttaatt aattaaatta aagaaacgac aaaagagtat 4260

gcaagaattt taaaacaact tagaggaata tgtatgagga tacaggctaa gctaccataa 4320

tgaagagacc tcgaaataca gtgagaagcg agacagaagt atctttcgtt ccatgtaaca 4380

ctccaggtggt tcagagcagc taagcagcta tgttccatag agtcattcag tgatccagat 4440

tattttcatc tgttgctctg ccattctcca ggatgttgtc cctataaaat tgtcaaagct 4500

cagtcagtgc caaacccatg tttcaacctt cagaaagtaa acgagtggtg gaaaacacat 4560

tcaatgtttt aaggccaaga ccttgaaaac tcactctctt agcctgaact tagattacat 4620

ggctgggccc acttaactat aggggaggct tggaaacata gtctctgaga agccatgtgt 4680

ccagctaatt ccctaatact aaagttgaaa gaaagaatgg attaaccagc agtataccac 4740

aaggtaacaa atgactagga ggatcaggct aggtggacta gaaaagagac agtcaattca 4800

gtgcaacaat tccatattga cactttccat gtagctgttg cttggctcta tctagagagg 4860

actcagaggt agtttagata aggcctttgc cctccaaata cagtctaagc agactgattt 4920

cctactggat gttcaacttt ggagtcttca gggatgagta gggcttctgt acgtggaaga 4980

gactatgagg gaacctgcac aggacaaggg tttgcataaa gacactgagg tagggacctc 5040

tcctgttgtg gggacagtga gaggcccagg tctccttgac tcacaaagtg cttactaagc 5100

acttactaga aattaagaag cagattataa tcaatatggg ttatccaatg tttggatgag 5160

caaggctcct tatctttct tcgttaatgt taatcacact cttttggatg gagacaaata 5220

tctgtggggg ctnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 5280

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 5340

nnnnnnnnnn nnggctcaga ctaaataatg tctaatctct tctccagtaa aacaatccgt 5400

ggttctcaga tagcactgtg ctggaggtag tggggtttga gggctgggaa gttgggagga 5460

ctgagcccct cccgctgagc agtttcgtcc agtttttcct gtaccagcct gtcatgttta 5520
```

38

```
ttccatgtga atgactccag aggcaaaatt caagcttttg aatagggcac aaattaactt 5580

gagtaccctt tcatttccct gtaggtgaac actcctctag ccctgccttt tgtcagtctg 5640

gagcccttgt tctaatctgt acacaccaga ggactttaca aggctttccc cagcctccag 5700

aattattctt ctgatccacc ctctactaaa ctcacccttt cctcagtgct aggacgttga 5760

aaaaccgaaa caaggcaaag ggccaattgt aataattcac actaaggcat gagtgactag 5820

gtttagtata ttaacactac ctaggatatt ctatttcttc caaaaggatc tcgttaatcc 5880

ttgaaattta acaactaatg gtatagattc taagcactgt gagtacttgt cagtggggga 5940

aagacatttt tgggctgaga gactttgcca ctgnnnnnnn nnnnnnnnnn nnnnnnnnnn 6000

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn aagaaagatg attagatnnn nnnnnnnnnn 6060

nnnnnnnnnn nnnaaaaaat aacatgagag nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6120

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6180

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6240

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6300

nnnnnncaca gagctagccg tgttggctgt cacccactca tgtgtgccagc ctgttggtct 6360

acctcttagt tgccatgtaa caggattctg gtgctttttcc tttgcccagg t cag atg  6417

gaa ccc agg acc cat cca act tgg ctg ctt cac att ttc atc ccc tcc   6465

tgc atc att gct ttc att ttc ata gcc aca gtg ata gcc cta aga aaa   6513

caa ctc tgt caa aag ctg tat tct tca aaa ggtaagtgag ttttattcat     6563

ggtaacccaa tgcactgggt gtctgcagca tgagccactg ctttgcactg caggcctatg 6623

gcttgctgct ttcatgctaa acccactcag agcttatgaa ccactttgag cttgtcttga 6683

tgattatttt tccccagaag aaaatggctc tcatcgtcag tgagctgaac ttcttacact 6743

gagtttttta aagggaatgt tttgttctta tgtctgaaag agtttgtctt attctttgag 6803

ccaagagctt tcatcagcct catgagagtg atgttattttt ggcaatgcag agagctacgt 6863

gctccgattt tgctggtggg aggttgccag gatcctttct gaggattcct tccattttca 6923

cccctctttt ccccagtctg gatatgacct gggtaaaacc caccccctct cccaggaatc 6983

tcaacctcac ggttggtaa ggaaaggaga aaggtttgtg aggccatttg gggatangga 7043

aacagctggt tggtggtgca ttaacgtctt tcagcagctc ccttcgagtt tctccttagc 7103

ctgttgtatt cttaccaaca cactcctgtt ctgttgtacc agctgggaca gagcatgctg 7163

aagcctttca gccctgattt cattgcttca ttgttcatgt gtctgtcttt ggtttcctgg 7223
```

39

```
gtggagcctg cccacaaaac ccccagaatg tatgcaggcc tagctggtgc tttcctaaac 7283

ggctcccttg tctgcactca atgaacttct ccaaagatct atacatggcc tcatctatag 7343

aaagagaaat gacatgtgga aataattcag taggagtttg cagcagcact atctgaggac 7403

taggggaatt ttaagtggtt gttatcttac atttatactc ataacttcta tattttcatc 7463

tgccataaaa tattgtcatg ttctatttgt ccattgccct atgtgtgtat gtattcactt 7523

gggtgctgac cacaatattt ctaactgtag aatgcaagga attgttgcca aannnnnnnn 7583

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7643

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnncttca tggtaaagga 7703

ggaggtatgt ccaacagaac ttcgactttt aaatagaacc acttcagaga gttgtgtcag 7763

gtgcacctca gttgtcttat cttctgccat tcttctttta cctctcacac ccatacctca 7823

gggttcaagg cctggggcct gaggactcct taataacttc agaaatgagc agctgagtgt 7883

tccgttccag ctgtctttgg gagaatggaa tggagtcaca ctcaaagata gagtggaaat 7943

aaatcctctc ctcatccttc accccaatct taagagtgag tgaggatatc agtagctccg 8003

agctgggagg taaagctcaa gttctaactg tgattaggag acctttctta caaataagaa 8063

ttaagtgaat aaatgtgcaa acaatttctt ttatattttt aatgaaccag agagaaatca 8123

tggttgccta tataaccctt gtctccaact cacttgcatt cagatctgct ttcttacatg 8183

tgtctgccat gcacacaaac ttgtgtgcca tggaaaaggg ttgagaactg ctggtgatgc 8243

agacagagct ttaaannnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8303

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8363

nnnnngcagc aagagaagga acattttaca gcttattggc cgaacttcac tgccgctagt 8423

gtggttcaac ttggactaca gagaaatctt cctaactggt ttccctgtat tcactcctgc 8483

tacctccaac ttggtctgtt ctcacttttt gctataatag gcttttaaaa atcataaatc 8543

taccatgtgt cctctgtcca gaccttctcc atggcttctt attgctcact ggatgaagtt 8603

ccaacgagcc caggatggtt tgactcatgt ctccagcttt aactgcatca ccatcacctt 8663

cattgtctaa agctctaacc acacaggatt ttctagtcct cagaggcatg gcagtctttc 8723

aattccgagt tttctcatac aatattgtct cttcttaaaa tatttttttct tgttgtccac 8783

ctgagttgga gtcatctttt aaatctcagc taagcttata cttcatcaag tctttcctaa 8843

ttctacctcc acgcaccaca cccattacat taaatcccct tattatatgt ttccatagca 8903
```

40

```
cctactttct tcttttcagt atactcagca cacaatcaca tgtctaggat ctgttttaat 8963

agcttggact accaattaaa ttgcatccct tttaattgtc cattgattcc tcaagtaccc 9023

acatgcccat cttagcaaga agttcagtgt ctccctctta tagcatgtac ttctccacct 9083

cccacaaact gccagaaagc ttacttagcc cacagggcca gtgctaggca gctaggttag 9143

tcctccagag ggccctggtt ttgagcagtt gctgtctact ccggccatgc agaatctctg 9203

gtccttccag atgtctccat ccactgtgca aaggtaacct tgctggttcc gatccccaca 9263

cagaccacag tgctacaaga ttacagttct tatggttccc caacacatgc tctgtcattg 9323

gtcccaaagc aggaccccta tgggttgatg aggtaggagg aggtccctgc cttagccaca 9383

gctgcacaca gccagcctct tcccttctag gccctcatgt tgagcctggg acgccagtcc 9443

taacttcctt ctcttcagtt cctcttaggg ccattggtat cctgaatttc ttagtccatt 9503

gcaaagttaa gtaaagaagc agcaggcttg gtccctttcc ttccagatgg cttcttagct 9563

cctgaacaga tttacccacc tatacctcag tgactagctc tgtgtactaa agtgtattgg 9623

gagggcagcc attattggtc cataaaaggt cctgcttacc attttcccct aagaggaacc 9683

attcaacagt ttggggctcg agggtgacct gctgggctct agagaagaag ctggcaactt 9743

ctgttgcaaa ataatgttaa attctgcttc atctgcttgt cttnnnnnnn nnnnnnnnnn 9803

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9863

nnnnnnnnnn nnnnnnnnnn nntaaacatt gaacctacta tatgcaggtg agtatgctag 9923

atnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9983

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10043

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10103

nnnnnnntgn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10163

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10223

nnnnnnnnnn nnnnnnnnga tttgaataca ggtctgtttg actccaaaac ttgtggccta 10283

tttgttgcaa aagtgcttaa tacaaattgg ttcagtcaat attattatct ttgaacaatg 10343

gaaggagaaa gtaagtttca atccaaaata attgagtgac ttatacattg acttgctgag 10403

ccaatggcaa agtcaagtta gaatccagca gaagtcacca gctacagaat ctagatcttt 10463

agaacatgtc ttcagatctt cagaacagtg tttcttaaac tctcttgtga aggaacagtt 10523

atcatcatag gctggtaaca gttcacctac cagcaccagc ccatgaacca gactctaagt 10583

ggcacagccc tagaagattg agccagaatt ttacagaggt ttaaagacca aatatgctgg 10643
```

41

```
tttatggtta cctgtggccc acagagaatg gcagcactaa cctcaggcat aaatgaggta 10703

cccactgaag ccaacattca agagcaattc ctatgggtta accattgggc tcctttcaaa 10763

tgcaaaccct catgaaagag actacagtgc tgaatagaga cctccaaatt ccaggccaag 10823

ctcaggatag tcatgaggga attactaaaa acctggtata tagggcaaaa gcagaattag 10883

gaatggactg atttcaggaa cccaggcaat ggcaggagtt gggcattaaa tcctaaaaga 10943

gaatcagagt gggagggaat atgtgaaatc agaggttaag aaaaaagtga aaacctnnnn 11003

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11063

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11123

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11183

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11243

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11303

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11363

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11423

nnnnnnnnnn nnnnnaaaaa attaaaagaa agatgtgaaa tcaaggaaac ttactggtga 11483

gcagcatccc attatgtgaa cttgtgcttc tgaaccagta acttgagtta ctttgagcca 11543

gtatcagtca cttatacctt agtgcaaaat taattgatca gacattctga cctggaccag 11603

ggaaggcagg cagaagtagc agtcaagact aaagcagaaa agggagagct aattctgcag 11663

ccagacattt cctggattga atacccaaat tagtccctca gcctttaagt gcctgagggc 11723

caggagtaga cagaggaatg gaaagtgtga gacttctttg ttcacactct ttgcctaggg 11783

gccagatttt gctttatgca ttaccatccg aagtcccagg ccacagtgaa catttgggct 11843

tcgctatgtg gatttattta gatttacttt ttgtcctgcc atattttaat ctataagcca 11903

aacagttttc tcattaatct tattccattt ctggaatttt tcctttttca gac aca aca 11961

aaa aga cct gtc acc aca aca aag agg gaa gtg aac agt gct          12003

gtgagtaagc atgattttta cttttcttc ttactttctt ttctctctca gcttgaattt 12063

taaagtaacc actgttctat taattcatgg aaggcaactg aatagttcca gcttatagaa 12123

tcttcctgtt tggtagcatt tcagcgaagc ctcgttctta gccccagaac aatcatgcca 12183

tcttttgctc ggtctatatt cctaagcact cctagatgat actgcactgg acctctggtc 12243

tcacatagtt agaaacagag ttaaaatcga acagcaaaga gaagatattc aactgcgatg 12303
```

42

```
caattgacaa tggatgtttt tgcaacaaac aatgattaag aagtacattg ttgtgggctc 12363

tgagtcaaga gtaatatggg aaaaacacaa gtctcttcat gaggttgaca ggtttggagc 12423

tggaatctgt ggaggaggaa ggatatgatc taggggtcag aagaagtggg ttactaaaat 12483

cattaagcct ggttggatga aaagcttaga ctcaggggaa gcagnnnnnn nnnnnnnnnn 12543

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 12603

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 12663

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 12723

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 12783

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 12843

nggcaccaag agatgagggg ggcagtcttg gccatatatt tggctgaagt aagtcaattt 12903

gtcattcctg catgagcctt tataaacaga agtaagtaac caactactat ttggtcattg 12963

gagttgtcca agaggccagg gttctgtcta atacctgttc atgcatgaac atgccaacct 13023

agattgcatg cagactacca gttttgggtt tttgtttagt tcagcaggat ttttctcagc 13083

tcactgcctc tcaaactttc agcaacaaaa ggacatctgt gatatcagaa tctaccactc 13143

taagtatttg gatgcaatag caatgaatat ctgagtaaat ctaggtgggg agtgggggca 13203

ccctgtagcc aaaatgattt aacaaaatca aaccaaaatt ttggaaatga tgccttggta 13263

caatgaagag actacttgag gtaggtttga cttatctaat atcttatttt ctttaccaat 13323

acctaatgag gaatttaaat atttctagat agctttggaa aggtccctta aagaggcacc 13383

agcataccac tgccagatct aatcccccca aacactgttt tcatcatcat catgtcatct 13443

cttgtctcta tagatcatat caaatccttc ccagagtttt tcaggccttt tgacaactag 13503

ccacatttca ctaagccaac tcatctacca ctcttcaaca aaacttttcc tcaagttgag 13563

ctgctccacc aacaccactg ccatgagctc attcccactt ctgtggcttt gctcatgttg 13623

gttatttttt tggagtgtcc tccctattcc ttcttacttg tcccaatccc aactttggc 13683

atggtctact ttaagataca gtaatgagta actttannnn nnnnnnnnnn nnnnnnnnnn 13743

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 13803

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 13863

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 13923

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn tttttaaaat ttaattcaan 13983

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 14043
```

43

```
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 14103
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 14163
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 14223
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nntacataca 14283
caacactctc ttgcttaatc tgtaagactc tctcccccac tcataccttt ttattttccc 14343
tctgcattgt acacacaatc tataccactc ttaagcacat gattacagcg ttattttctg 14403
gctgcttcta tgtgtctata ttttaggtcc acctggtcaa tataataaag tgggatatta 14463
gtgttaatgc aactatatgg tatttgatat ttgtctttct gtccgtttat caatgtttct 14523
tatagnnnnn nnnnnnnnnn nnnnnnnngg tgtctgattt tgaccaaatt tgactaaata 14583
cnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnngt tgagtctatc aactcaaaaa 14643
agaataacct acaacaataa caagtttcag aacatttttt aaattactga ttttatgagn 14703
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnngaat ccattgccta 14763
gaaattgcca tggttaagat tttaatattg ctcaggccca gacagctcag ggctttgaca 14823
ttcccacacc cattctctgc catcccagtt ctatctcatc ccaaaaccat ccattatgag 14883
gagagtgtac agctctaggc tgcccgggag ccatcccgca ctctcatttt gtgactcggc 14943
atcttgggag atggagtctt gggacttagc ctggacatgt cccttgcatg tacttcttac 15003
aagactttta ttcagatgaa tattttccct tccaacttaa gaagcacagg gcttgctggt 15063
tttgcttcac taaccagcaa ctgaagcaag acctgacttg tgaaaatgcc taatagagtt 15123
cagtattagc gctgtttcac catttcctgg atctcttgcc tttgtgcaca tgatagaatt 15183
gcacttctct gtgattaatt tgtagttaag tgtggtcatg taactcgctt tggtcaatta 15243
aatgtaagca taagtgatgc gtgttatttc tgggtagaag atgtaagagt tggcatatgc 15303
tttgccatat tttctttatc catctggcat ggtaaccagt aacattctag gtagtaattg 15363
ctccatcagt ctcagtctct gagtgactaa aattgacaga gtcccctgct gaccctcaat 15423
gtacatggaa catgaacaag aataagcttt tgtttttat attgagattt tggagttgtt 15483
tgttcctaca gcattaccta gtttactcta atacaacatg gaaaaaactg gaacctataa 15543
taaatagacc ctacgttgcc atttaaactt ctagttctga ggataataa tgtgggaaa 15603
tactttctat ataataaaaa aatagaaaat tgcaaaataa aaatatactt atgtatcatt 15663
catgtcctat taaaaatgtt atttatagac tcaccatatt cccttcctcc agaaaaatag 15723
```

44

```
aagtaaaaat atgaaaatgc ctgtaatcat gtttttggat tatggaatca agtattgctt 15783

tttactttta tgttttctga atttttgttg tacttcacta cattttgag tgccctgatg 15843

tattactttc aaaaagaaga agaatacttt ctgaagccat ttcaaccatc cccactcacc 15903

tctctagatc ccagtaacca aatacattat ataggactct tcatcagtcc ttatcaagtt 15963

taggaagggc gatgctatac cttctttaaa ggacacctac caatgtctta gttgcctttc 16023

aaagactcct agcacagcta aatgtgatgg atatgctcta aggatataag agctgaagtg 16083

acttgcataa ggtcatatca taacttactg ttagaaatgg agctagaact cagacccact 16143

gagtccttgt ctgtgacaca ctgccctttc catttgtgga agttgttctt gtatctaact 16203

ttatctgtgc tactatttgg gcctagccat tctccctctt atgcagacaa gcagataaac 16263

agtaaaactt taggagtgga ttatgatacc atagatatat atcatctatc ctttacaaaa 16323

tagttattac agtcatcaag ccttggttag agtttacaga ccatgtatcc tagctannnn 16383

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnntgaaaca 16443

gtagcagacc aaaagaagtc atgattccca gcatagtgct nnnnnnnnnn nnnnnnnnnn 16503

nnnnnnnnnn nnnnaaacct gtattaagtt tctgttattt gcaagaccct gtcagttagg 16563

ccatgtggga actagaagga tgaatttatc agtcatccaa gattcttaca attaagtatt 16623

accgataagg tactcaagaa acagttctca ttcacataat ttgggttaaa acaaaaagaa 16683

gccagctttc tatatacttt tggtccagtc tttacgtttt ttgtttttgtt ttgtttgttt 16743

tcatgagtat cccgacttcc ttctaagaac ttccacctga gaactgacca cagcgtcagc 16803

attccacatg ggtgtgtttc ctttccccctt tcccatttca gtggtttcca atttctttt 16863

cttttggcac tataaacctt tcgcaaagga aatattagac agaactccta catgtcaagc 16923

aaattaaaat agtggtgaaa ttagagtgga ggacataatc accctatcat ataggctatt 16983

tgtccatatc atatttgtcc ctacaaaggc ctctaaggnn nnnnnnnnnn nnnnnnnnnn 17043

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 17103

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 17163

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 17223

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 17283

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 17343

nnnnnnnnnn nnnnnnnnnn nnnctctaag ataccttgtg cttcttggaa catatttgga 17403

aaatcatgta gctctcaaat tatccctatg tcctgaagcc caccttacca tcaatcctca 17463
```

```
gaaatacaca acccatgtca ggcaacttca cactttcttt cttcaggcag cacagttgtc 17523

tcagggaggg aggagagtgc tattagcaag aggagtcact aacagcttca ctcacctgtg 17583

cccatgaatt tttaatggtg tgaaaagtct gtgtattttt tatggttctc tatggctcat 17643

ataagaggga caaacaatac atgaaagttc agagatgggg acaaatatca ctttaagctg 17703

gggataatca gggaagaaga aagtttttcat ggagaaggtg acttttgaat caggaagaaa 17763

tggaaacagc aaactcttct agatagagga gacactaaca ggaaaggcag agaggcagga 17823

aggtgtggga aagtgcacgt gaccccggttc agagagaaag ccaggtgtga gataaggggg 17883

aaagactgtt agggcatgta ttgtaaacca ctaattccag gcaaaagtta gattttactt 17943

actaagcaag agtgcttcag ttagatccta gcaggaaatg gagggtatgc ttagaagagg 18003

taactgaggc aagtttaatt tataaaggtg tgtgcagcat taagggaaac cagcaaggga 18063

tactgagcat gccaggatgc aagagcaggc agggaaggtg actattccta ggtctgaagg 18123

agaaagggga gggagcagtt cccagaaccc tagtaaaaat ggcaatgaga aagtccatc 18183

tggcaggacc tatggtcttt aacagaggga caaagtcaac ccacaacttg tctgggaggt 18243

tgctgaggaa tagataccc aacctctctc tcaacccact gcaacactct ttttccccta 18303

gactgagccc agtcaaagac agaggggagga gcccagtgat gcagtctgca atgtcatcat 18363

cctggagcat gaatagagtg cagcaggtg aataatgagt ctgcaggaat taatagaaat 18423

atctgacaca atagggaact ataagaggtt ttgaatagga gaggcccctg aaatgtgctc 18483

caatattact gaactatgtg tggcccaaag aatggaagag gaacagctct tgcaataggt 18543

ctgaggagag aagctgaaga cttggactag ggcaatggta aaaactgtgg aaagaagttt 18603

taaatgaaaa gttttaaacc atgcggcttc cagctagatg aacttttta aaaaaattag 18663

ttcctcactc aaattttggg gaggttatat attttctaat cataaaaat gattttcctt 18723

atttgtggggc ttttctcccc ag atc tga acctgtggtc ttgggagcca gggtgacctg 18781

atatgacatc taaagaagct tctggactct gaacaagaat tcggtggcct gcagagcttg 18841

ccatttgcac ttttcaaatg cctttggatg acccagcact ttaatctgaa acctgcaaca 18901

agactagcca acacctggcc atgaaacttg ccccttcact gatctggact cacctctgga 18961

gcctatggct ttaagcaagc actactgcac tttacagaat tacccactg gatcctggac 19021

ccacagaatt ccttcaggat ccttcttgct gccagactga aagcaaaagg aattatttcc 19081

cctcaagttt tctaagtgat ttccaaaagc agaggtgtgt ggaaatttcc agtaacagaa 19141
```

46

```
acagatgggt tgccaataga gttatttttt atctatagct tcctctgggt actagaagag 19201

gctattgaga ctatga                                                  19217
```

<210> 19
<211> 18
<212> PRT
<213> Homo sapiens

<400> 19

```
Met Ile Phe Leu Leu Met Leu Ser Leu Glu Leu Gln Leu His Gln
  1               5                  10                  15

Ile Ala
```

<210> 20
<211> 102
<212> PRT
<213> Homo sapiens

<400> 20

```
Ala Leu Phe Thr Val Thr Val Pro Lys Glu Leu Tyr Ile Ile Glu His
  1               5                  10                  15

Gly Ser Asn Val Thr Leu Glu Cys Asn Phe Asp Thr Gly Ser His Val
                 20                  25                  30

Asn Leu Gly Ala Ile Thr Ala Ser Leu Gln Lys Val Glu Asn Asp Thr
             35                  40                  45

Ser Pro His Arg Glu Arg Ala Thr Leu Leu Glu Glu Gln Leu Pro Leu
         50                  55                  60

Gly Lys Ala Ser Phe His Ile Pro Gln Val Gln Val Arg Asp Glu Gly
 65                  70                  75                  80

Gln Tyr Gln Cys Ile Ile Ile Tyr Gly Val Ala Trp Asp Tyr Lys Tyr
                     85                  90                  95

Leu Thr Leu Lys Val Lys
                100
```

<210> 21
<211> 91
<212> PRT
<213> Homo sapiens

<400> 21

```
Ala Ser Tyr Arg Lys Ile Asn Thr His Ile Leu Lys Val Pro Glu Thr
  1               5                  10                  15

Asp Glu Val Glu Leu Thr Cys Gln Ala Thr Gly Tyr Pro Leu Ala Glu
                 20                  25                  30
```

Val Ser Trp Pro Asn Val Ser Val Pro Ala Asn Thr Ser His Ser Arg
             35   40

Thr Pro Glu Gly Leu Tyr Gln Val Thr Ser Val Leu Arg Leu Lys Pro
    50   55   60

Pro Pro Gly Arg Asn Phe Ser Cys Val Phe Trp Asn Thr His Val Arg
65   70   75   80

Glu Leu Thr Leu Ala Ser Ile Asp Leu Gln Ser
                85   90

<210> 22
<211> 44
<212> PRT
<213> Homo sapiens

<400> 22
Gln Met Glu Pro Arg Thr His Pro Thr Trp Leu Leu His Ile Phe Ile
1   5   10   15

Pro Ser Cys Ile Ile Ala Phe Ile Phe Ile Ala Thr Val Ile Ala Leu
             20   25   30

Arg Lys Gln Leu Cys Gln Lys Leu Tyr Ser Ser Lys
          35   40

<210> 23
<211> 17
<212> PRT
<213> Homo sapiens

<400> 23
Asp Thr Thr Lys Arg Pro Val Thr Thr Thr Lys Arg Glu Val Asn Ser
1   5   10   15

Ala

<210> 24
<211> 11
<212> PRT
<213> Human immunodeficiency virus type 1

<400> 24
Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg
1   5   10

<210> 25
<211> 15
<212> PRT
<213> Artificial Sequence

48

```
<220>
<223> Description of Artificial Sequence: Internalizing
      domain derived from HIV tat protein

<400> 25
Gly Gly Gly Gly Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg
 1               5                  10                  15


<210> 26
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
      Oligonucleotide Primer 2515-27

<400> 26
cataatagag catggcagca atgtgac                                        27


<210> 27
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
      Oligonucleotide Primer 2524-63

<400> 27
gggtcctgga gtggctggtg ttg                                           23


<210> 28
<211> 14
<212> PRT
<213> Homo sapiens

<400> 28
Leu Phe Thr Val Thr Val Pro Lys Glu Leu Tyr Ile Ile Glu
 1               5                  10


<210> 29
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
      Oligonucleotide Primer 2515-24

<400> 29
gtggctcttt cacggtgtgg ggatg                                         25
```

```
<210> 30
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
      Oligonucleotide Primer 2538-68

<400> 30
ccagtgtcaa agttgcattc cagggt                                    26
```